Section 6
Evaluation of Sources of Contamination in the IRW

6.1 Purpose

The purpose of this section is to evaluate and document a link (if any) between poultry land waste disposal and environmental contamination in the IRW. This objective can also be stated as a scientific hypothesis:

- Land application of poultry waste affects the chemical and bacterial water and sediment composition of the IRW and the affect is observable in surface water, groundwater and sediments collected from the IRW.

A second objective was to evaluate other potential sources of the environmental contamination in the IRW. The sampling schemes and evaluations were designed so that all major sources of contamination would be identified. However, specific other sources evaluated were cattle manure deposited in fields and rivers and waste water treatment plant (WWTP) direct discharge into IRW rivers. This objective can also be stated as a scientific hypothesis:

- Cattle manure deposited in fields and rivers and WWTP discharges into rivers affect the chemical and bacterial water and sediment composition of the IRW and the affect is observable in surface water, groundwater and sediments collected from the IRW.

As discussed in Section 3, a conceptual site model (CSM) is a part of the process of systematic planning process. The CSM is a functional description of the contamination problem. The CSM is initiated at the start of the project and updated throughout the project as information is developed and evaluated. The CSM is usually accompanied by a diagram **(see Figure 6.1-1)** which graphically illustrates the relationships among:

- Locations of contaminant/waste sources and locations where contamination exists
- Potentially contaminated media and migration pathways
- Potential human and ecological receptors

The geology, hydrogeology and migration pathways are discussed in more detail in Sections 6.3 and 6.5.

6.2 Evaluation Approach

The overall evaluation was conducted using multiple evaluations and investigations or multiple lines of evidence. The results of the multiple evaluations and investigations were then used to determine overall conclusions concerning the poultry waste in the subsurface. Furthermore, two of the samples that contained quantifiable concentrations of the PLB were base flow samples from the small tributarties, which consist mainly of groundwater. In summary, the PLB has been detected in all environmental components of the IRW.

6.11 Chemical and Bacterial Signatures using PCA Techniques

6.11.1 Introduction

Principal component analyses (PCA) is a multivariate statistical technique. Multivariate means that multiple response variables or parameters (contaminants) were measured and are available for evaluation. Multivariate analyses make use of correlations between variables to help explain important relationships and reduce the number of variables needed to describe relationships of concern. In an environmental context, PCA is used on sites with a large number of contaminants and allows us to determine the differences and relationships among all of the contaminants. These relationships are used to identify sources of contamination.

PCA is probably the oldest and best know of the techniques of multivariate analysis (Jollifee, 2002). PCA was first introduced by Pearson (1901) and developed independently by Hotelling (1933). PCA is a well established statistical technique in environmental and other scientific disciplines. In particular, PCA has been used to evaluate sources of contamination in the environment. A list of publications concerning use of PCA in evaluating environmental conditions is provided in **Table 6.11-1**. **Table 6.11-2** provides excerpts from some of the publications with respect to source evaluations and conclusions. These documents show that PCA is a reliable and accepted scientific techniques and can be used to identify sources of contamination in watersheds such as the IRW.

Table 6.11-3 provides a comparison of important attributes of four selected publication studies used to identify sources of contamination in watersheds.. The same attributes for the IRW evaluation are also provided in the table. As shown in **Table 6.11-3**, the IRW study and PCA evaluation were very similar to the published studies in terms of key study attributes for PCA such as watershed size, number of samples, number of parameters used in the PCA, number of data points, number of sources identified, and percent variance explained by meaningful principal components. The IRW discussed in this report has a significant more number of sampling stations and the major principal component explains more variance than the other studies summarized in **Table 6.11-3.**

6.11.2 Steps of PCA

The steps used to perform the IRW PCA were as follows:

- Step 1: Formulate a conceptual site model
- Step 2: Define objectives and state hypothesis

- Step 3: Prepare a comprehensive list of parameters to be analyzed

- Step 4: Design a systematic sampling program for all environmental component sin the conceptual site model. The elements of a systematic sampling plan include identification of the objectives (step 2 above), identification of the data needed (step 3 above - parameters to be analyzed), description of the intended data use, formulation of the sampling approach and scheme (sampling locations) and determination of sampling methods (standard operating procedures – SOPs).

- Step 5: Collect samples and conduct analyses according to the sampling program design and SOPs.

- Step 6: Compile a "clean" database. This includes field and laboratory data input, transcription checking, data quality review, addition of data qualifiers, assignment of groups and resolution of any location conflicts. This step also includes creation of subdatabases from the main database (e.g, a database with surface waters data only).

- Step 7: Perform distributional and data exploration analyses. This includes generation and evaluation of descriptive statistics (count, minimum, maximum, mean, median, variance and standard deviation), probability plots and data distribution, number of censored (nondetect) cases, data transformations (logarithmic, square, etc), Pearson correlation matrix, outlier identification and overall evaluation of parameter variance. Graphical displays include correlation plots and probability plots.

- Step 8: Identify parameters that meet PCA criteria. PCA criteria include maintaining a large number of parameters, selecting parameters showing good variability, selecting parameters with substantial amounts of detections, selecting parameters that were consistently analyzed and eliminating redundant parameters. This step was performed interactively with step 7.

- Step 9: Normalize and standardize data and perform statistical evaluation of parameter correlations and relationships. This is the formal PCA evaluations process that includes computation of principal components using various rotations, computation of parameter coefficients for each principle component parameter for each rotation type, determination of variance explained for each principal component for each rotation type, standardization of the concentrations (using z statistic), and graphical displays of parameters coefficient magnitude, variance explained, parameter relationships and principal component vs principal component scatter plots.

- Step 10: Identify major principal components that explain the chemical and bacteria composition variability observed in the IRW surface water, groundwaters and sediments. This includes determination of number of principal components to evaluate and is performed interactively with step 9.

- Step 11: Calculate principal components scores for each major principal component for each individual sample. To calculate a principal component score for each individual sample, the parameter coefficient is multiplied by the normalized parameter concentration. This is done for all parameters and the products are summed to yield one value for each sample for each principal component. For example, one sample will have both a PC1 score and a PC2 score (if two major principal components are identified).

- Step 12: Evaluate whether the major principal components are associated with specific sources. This step consists of two evaluations: 1) comparison of the principal component parameters to the composition of known waste sources and 2) a spatial and temporary analysis of individual principal component scores (for all major principal components). The spatial/temporal evaluation evaluates principal component scores in relation to the location of the sample (distance from sources), group or environmental component (e.g, edge of field), sample conditions (e.g., high flow, base flow), poultry house density, and reference locations.

- Step 13: Use the PC scores to determine the samples and locations in the IRW that are impacted by the major sources of contamination. This evaluation includes comparison of principal component scores for reference locations and areas with minimal potential contamination (e.g., locations with low poultry house density). This step is performed in conjunction with step 12 spatial evaluation. Step 12 evaluates specific known locations; step 13 is more of a basin wide evaluations. However step 13 also evaluates specific contamination sources.

- Step 14: Perform investigative and sensitivity analyses. Sensitivity analyses were performed on the number of parameters, specific parameters (e.g., arsenic and nickel), the groups or types of samples from environmental components (e.g., combinations of different environmental components), types of analyses (e.g., various forms and analytical methods for phosphorus) and specific samples (e.g., outliers).

- Step 15: State and document conclusions concerning source identification, dominance of sources, locations of contamination and robustness of analyses.

Each of the steps and the resulting conclusions are discussed in the following paragraphs:

Step 1: Formulate Conceptual Site Model

As previously stated, the conceptual site model was formulated in conjunction with Dr. Bert Fisher (Fisher, 2008). The site conceptual model was previously discussed in section 6.1 and shown graphically in **Figure 6.1-1**. The site conceptual model identified all major fate and transport components in the IRW.

Step 2: Define Objectives and State Hypothesis

Each sample collection and analyses program had defined objectives. The objectives of the each of the sampling programs are provided and discussed in Section 2. Section

2 includes the identification of project objectives, the type of data needed, the intended data use, the sampling design and sample collection procedures.

The overall objective and hypothesis of this Section 6.1.1 (Identification of Sources of Contamination in the IRW) are stated at the beginning of this section (Section 6.1).

Step 3: Prepare a Comprehensive List of Parameters to be Analyzed

The process for selection of analytical parameters was discussed in Section 3.2. This list of parameters is also compared to the chemical and bacterial characteristics of potential major waste sources in the basin (poultry waste, cattle waste and WWTP discharge) in Section 6.4.1. As discussed in these sections, the extensive list of parameters consists of all major constituents of the potential waste sources and constituents that can be used to distinguish among the sources (i.e., constituents that are different in the various wastes).

Step 4: Design a Systematic Sampling Program

The designs of the sampling programs for each environmental component are discussed in Section 2. Each program was designed to achieve the stated objective and maximize the probability that representative samples were collected.

Step 5: Collection and Analyses of the Samples

Sample collection using standard operating procedures is discussed in Darren Brown's expert report. An evaluation of the sampling procedures related to the intended use of the data is also provided in Section 2 of this report. The discussion of the analyses of the samples including QA/QC and data useability is provided in Section 3 of this report. As discussed in these sections, the sampling design, and collection and analyses of samples resulted in data (except for rejected data) that were representative, precise, accurate, comparable, and complete and could be used for their intended purposes including PCA evaluation of sources.

Step 6: Preparing Data for PCA

The compilation of the main ACCESS database was discussed in Section 4. This includes transcription and data transfer checks and addition of data qualifiers. The quality review process and assignment of data qualifiers is discussed in Sections 3.6 and 3.7. The main database is named IllinoisMaster.mdb.

After the main database (IllinoisMaster.mdb) was finalized, data were extracted from it by queries and placed into two separate EXCEL workbooks, referred to as the main PCA databases, one for water data and one for solids data:

- PCA_Main_Database_Water.xls
- PCA_Main_Database_Solids.xls

The main PCA databases were used to further evaluate, process, and compile the data, and to develop, apply, and document various protocols for data use in the PCA. This was accomplished by dividing each main database into separate named worksheets. The worksheets in the main PCA water database (PCA_Main_Database_Water.xls) include the following:

- Water (SWGW)

  This worksheet in the main PCA water database contains the portion of the data selected or retained for PCA use. The process used to retain data for inclusion in this worksheet is discussed in more detail below and in Step 8. The documentation of this process is contained in this and other worksheets in the main PCA water database as discussed below.

  This worksheet contains selected or retained original data records (in linear or database form), with all of their original database fields, as they were extracted from the main project database (IllinoisMaster.mdb). In addition to the original data records, new or created records were added to the data in this worksheet. These include: (1) records for the three phosphorus forms (P_Sol_Reac, P_T, and P_TD) based on an established phosphorus protocol and (2) records used to set total concentrations to dissolved concentrations for geoprobe samples. These protocols are discussed further below and in Step 8. A field named: ConversionNote was added to the worksheet in order to document how these new records were created. In all cases where original records were used to create new records, the original records are always retained in this worksheet; however, the original variable name is indicated with an "x" prefix in order to distinguish it from the other variables. This process was implemented to establish complete documentation and to allow tracking of these changes.

  In addition to the ConversionNote field, nine other new fields were added to this worksheet. These nine fields are all required by the CDM-developed EXCEL Add-In program: EDAnalyzer (described in step 9) and are as follows:

  - EDA_Group – Identifiers used to divide the samples into logical groups, based on type of sample, location, flow conditions, etc. This field is used by EDAnalyzer to select appropriate groups for investigative and sensitivity analyses.

  - EDA_Sample – A unique sample identifier assigned to each sample, which includes the location identifier (or station where the sample was collected), the date of sample collection, the type of sample (e.g., "SW" for "surface water"), the depth interval, and other pertinent information about the sample. Duplicate samples (e.g., field splits) are given the same EDA_Sample identifier as the original sample. EDAnalyzer handles field duplicates by averaging the data (i.e., duplicate sample data are averaged prior to PCA).

  - EDA_Location – An identifier of the location (or station) where the sample was collected. This is also included in the EDA_Sample identifier (see above). Multiple samples collected over time (i.e., during the course of this investigation) at the same location or station will have the same EDA_Location identifier. EDAnalyzer has an option that allows averaging of data by EDA_Location instead of by EDA_Sample. This option was investigated during previous, preliminary PCA runs but was not used for any of the current runs. Instead of averaging data prior to PCA, any averaging of PC scores by location was conducted following PCA, as applicable.

- EDA_Variable - The assigned parameter (or variable) code name for processing by EDAnalyzer and SYSTAT, restricted to a maximum length of 12 characters. The process used to assign these code names to the variables is documented in the worksheet: Water (Variables) and is described further below.
- EDA_Value - The analytical or result value. This is a copy of the values contained in the original Value field as they were extracted from the main project database. The EDA_Value for censored data is set to the censoring limit (e.g., analytical detection limit) in the main project database.
- EDA_ValOp - A qualifier assigned to each data value. This is a copy of the original ValOp field as it was extracted from the main project database, which contains one of the three qualifiers: "<" for left-censored or nondetect data, "=" for quantified data, or ">" for right-censored data. For purposes of processing in EDAnalyzer, all "<" qualifiers where replaced with "U" qualifiers. EDAnalyzer recognizes "U" qualified data as meaning left-censored (nondetect) data and applies a multiplier to the EDA_Value as a selected option. In all of the current PCA runs the multiplier was set to 0.5, meaning that the EDA_Value was set to one-half of the detection limit.
- EDA_UnitsID - The units identifier assigned to the variables. The process used to assign the units identifiers is documented on the Water (Variables) worksheet in the main PCA water database (described further below).
- EDA_Y and EDA_X. The Y and X coordinates for the sample locations or stations. This is a copy of the original Latitude and Longitude fields as extracted from the main project database.

■ Water (Out)

This worksheet in the main PCA water database contains the portion of the data not selected or retained for PCA use. The decision to not retain these data was based on record counts (completeness), percentages of left-censored (nondetect) data, and previous, preliminary PCA runs.

This worksheet contains original data records (in linear or database form), with all of their original database fields, as they were extracted from the main project database (IllinoisMaster.mdb). A field named: ReasonNote was added to the worksheet in order to document the reason for not retaining or not including the data for purposes of the PCA. Reasons included: inaccurate data (e.g., phosphorus by method 6010), insufficient number of analyses, insufficient number of detections, no value reported (e.g., due to insufficient sample size), not applicable to PCA (e.g., barometric pressure), redundant variable (e.g., carbonate, which is redundant with alkalinity), and rejected data. All of these reasons were used to identify data that did not meet PCA criteria, as further discussed in Step 8.

- Water (Variables)

  This worksheet in the main PCA water database consists of a list of the original fields: ParamID and UnitsID for all retained variables in the Water (SWGW) worksheet. The EDA_Variable and EDA_UnitsID fields located adjacent to the original fields show the assignments made for purposes of processing in EDAnalyzer and SYSTAT. This worksheet documents the variables (EDA_Variable) and units (EDA_UnitsID) assigned to the data for use in the PCA.

- Water (P Protocol)

  This worksheet in the main PCA water database contains a cross-tabulation of the various phosphorus data (in rectangular or tabular form) retained for PCA use and documents the protocol for assigning data to the three forms of phosphorus used in the PCA: soluble reactive phosphorus (P_Sol_Reac), total phosphorus (P_T), and total dissolved phosphorus (P_TD).

- USGS (N DB)

  This worksheet in the main PCA water database contains a copy of the USGS data for various nitrogen analyses. This worksheet was used to construct a cross-tabulation of these data, provided in the worksheet: USGS (N CT) discussed below.

- USGS (N CT)

  This worksheet in the main PCA water database contains a cross-tabulation for use in evaluating and assigning the USGS nitrogen data to the appropriate variables. The original ParamID: Ammonia Nitrogen refers to USGS method code P00625, Ammonia + Organic Nitrogen (mg/L as N). Since this is the same as total Kjeldahl nitrogen (TKN), data corresponding to P00625 were assigned the code: TKN in the EDA_Variable field.

The worksheets in the main PCA solids database were similar to those in the water database and are therefore only briefly summarized below:

- Solids (SD)

  This worksheet in the main PCA solids database contains the portion of the data (in linear data records or database form) to be retained for PCA use.

- Solids (Out)

  This worksheet in the main PCA solids database contains the portion of the data not retained for PCA use.

- Solids (Variables)

  This worksheet in the main PCA solids database contains a list that documents the variables and units assigned to the data used in the PCA.

- Solids (P Protocol)

This worksheet in the main PCA solids database contains a cross-tabulation of the various phosphorus data (in rectangular or tabular form) retained for PCA use and documents the protocol used to assign data to the form of phosphorus used in the PCA: total phosphorus (P_T).

The process followed in retaining (or not retaining) data for PCA use and applying the various protocols documented in the two main databases was developed based on Steps 1-5 and the experience gained during a previous, preliminary set of PCA runs. As this preliminary work was conducted on an incomplete database (recently collected data were not included) they are not discussed further or presented in this report.

In summary, for the water samples, a total of 82,111 individual data records were extracted from the master database or created during processing in the main PCA water database. Of these, 49,088 records were retained for use in the PCA and 33,023 records were not retained. The retained data contained results for 66 analytical parameters, which were each assigned one of 40 unique variable codes for use in the various investigative and sensitivity PCA runs described in this report.

Similarly, for the solids samples, a total of 18,546 individual data records were extracted from the master database or created during processing in the main PCA water database. Of these, 13,101 records were retained for use in the PCA and 5,445 records were not retained. The retained data contained results for 98 analytical parameters (note: this number is higher than in the case of the water database due to inclusion of both dry weight and wet weight data), which were each assigned one of 41 unique variable codes for use in the various investigative and sensitivity PCA runs described in this report.

Individual EXCEL sub-database files were created from the main databases for use in the actual PCA runs; i.e., for import into EDAnalyzer. These sub-database files were given names all beginning with "Subdatabase" and include a sequence number that indicates the date (month and day) of creation. The date indicator was used for documentation purposes, in order to allow tracking of the various PCA runs and result files to a particular sub-database. The sub-databases were exact copies (on the date indicated) of the data contained in the nine EDA fields located in the retained data worksheets of the two main database workbooks. Following is a listing of the sub-database files used in the PCA:

- Subdatabase_Water_0427.xls
- Subdatabase_Water_0428.xls
- Subdatabase_Solids_0429.xls
- Subdatabase_Solids_0430.xls
- Subdatabase_Solids_0501.xls
- Subdatabase_Solids_0502.xls

Step 7: Perform Distributional and Data Exploration Analyses

Data exploration or exploratory data analysis (EDA) is a key component of, and is integrated directly into, the PCA conducted during this investigation. In fact, the name of the CDM-developed EXCEL Add-In program: EDAnalyzer means "Exploratory Data Analyzer". EDAnalyzer is a tool specifically developed for analysis of multivariate datasets, allowing interactive EDA in order to: (1) examine the distributions of and select appropriate variables (analytes) for PCA, (2) determine appropriate variable transformations, and (3) identify possible outliers for further review and/or elimination. In addition, EDAnalyzer performs PCA (via a shell to the SYSTAT program) and loads, displays, and saves PCA results for further examination.

EDAnalyzer was not the only approach used for EDA in this investigation: other EDA methods were conducted outside of the EDAnalyzer program. The results of these other methods are discussed in appropriate locations in this report.

EDAnalyzer operates by first loading the appropriate sub-database file (listed at the end of Step 6). Selections are then made of the various groups (EDA_Group), variables (EDA_Variable), and samples (EDA_Sample) of interest to a particular analysis or run. An option under sample selection is used to set the criterion to be used to limit the retaining of samples to a desired level of completeness of the variables, e.g., samples with data for at least 20 of 26 variables. Another option is used to set the multiplier for handling nondetect data (note: for all PCA runs conducted during this investigation, the multiplier was set to 0.5, meaning that the result was set to one-half of the detection limit). The program then generates a cross-tabulation of the data (samples in rows by variables in columns) based on the selections and options. During generation of the cross-tabulation, the nondetect multiplier is applied and the results for replicates (e.g., field splits) are averaged. The program then generates descriptive statistics and a correlation matrix for the cross-tabulation.

The correlation matrix was used only as a means of identifying possible "holes" in the matrix for purposes of the PCA, and was not used as input to the actual PCA. Holes in the correlation matrix are due to variables with an insufficient number of results relative to other variables. These variables were identified during previous, preliminary PCA runs and used to remove variables; therefore, for the current PCA it was typically not necessary to examine the correlation matrix for holes that would prevent the PCA from running.

The descriptive statistics generated for each variable were as follows:

- Count
- Mean
- Median
- Minimum

- Maximum
- Standard Deviation
- Skewness
- Kurtosis

In conjunction with the descriptive statistics (listed above), probability plots (or pplots) are generated in order to examine the distributional shape of the data for each variable. An interactive tool is used to examine the effect of various transformations on the distributions. The possible transformations available in EDAnalyzer are: natural logarithm, base-10 logarithm, square, and square-root. This step is important in the PCA for two reasons: (1) it is desirable to have distributions that are near-normally shaped and (2) it is desirable to re-scale the data so as to minimize the affect on the PCA of variables with widely varying concentrations, distributions, and units of measure. In practice, for most of the PCA runs, data were base-10 log transformed for all variables (although there were exceptions) to obtain near-normal distributions for most of the paramets and to minimize the affect of highly variable concentrations and units of measure. This is a common practice for environmental data which are typically log-normally distributed. As an example, the probability plots for run surface samples (SW3) are provided in **Appendix E.**

The descriptive statistics and pplots were also used to identify anomalous data or outliers. Such outliers were always checked to verify that they were not the result of transcription errors in the project database or on laboratory reports. In cases where transcription errors were identified, these were corrected in the main database and a new sub-database generated for PCA (note: this iterative process is one reason for the multiple sub-databases listed at the end of Step 6: to allow documentation of these corrections). In cases where transcription errors could not be verified for the outliers, they were either retained in the PCA or were eliminated by removing an entire sample. Such eliminated samples (which were always few in number) were removed via an interactive tool on the generated cross-tabulation. The following samples were removed as outliers in selected and corresponding PCA runs:

- EOF-SPREAD073B:6/18/2006:SW:S:-:-

This is an edge-of-field runoff sample that exhibited anomalously high concentrations for several variables. Some of the values reported seem to be laboratory errors; however, the laboratory error could not be confirmed.

- LK-01:5/17/2006:SW:S:0:-
- LK-02:5/16/2006:SW:S:0:-

These are surface water samples collected from Lake Tenkiller that exhibited anomalously high sulfate values (7,055 and 7,032 mg/L, respectively). These values are obvious laboratory errors.

- MAN-BC-20D:3/31/2008:SW:S:-:(SPLP-4-1)
- MAN-BC-22F:4/1/2008:SW:S:-:(SPLP-4-1)
- MAN-BC-24D:4/3/2008:SW:S:-:(SPLP-4-1)
- MAN-BC-24F:4/3/2008:SW:S:-:(SPLP-4-1)

These are cow manure leachate samples that exhibit extremely high concentrations for several variables. All 4:1 leachate samples were excluded from the PCA in lieu of 20:1 leachates which are considered more realistic of runoff.

- FAC-16:12/14/2007:SW:S:-:(SPLP-4-1)
- FAC-16:12/14/2007:SW:S:-:(SPLP-20-1)
- FAC-17:12/19/2007:SW:S:-:(SPLP-20-1)

These are chicken waste leachate samples that exhibit extremely high concentrations for several variables.

- EOF-Q1:6/17/2006:SW:S:-:-
- EOF-Q2:6/17/2006:SW:S:-:-
- EOF-Q3:6/18/2006:SW:S:-:-
- EOF-Q4:6/18/2006:SW:S:-:-

These are edge-of-field samples that were not selected because the actual locations and collection process could not be documented.

In summary, the EDA (descriptive statistics and the pplots) were used to help identify a set of variables and samples to be retained for the PCA. This process is discussed in further detail in Step 8.

Step 8: Identify Parameters that Meet PCA Criteria

The identification of parameters (variables) that meet PCA criteria was an iterative process. Ultimately, this determination was made during EDA as discussed in Step 7. However, much of the actual identification and selection occurred and is documented in the main databases (Step 6) based on previous, preliminary PCA runs and other calculations. Overall, the criteria used to identify parameters for PCA are stated and discussed below:

- Include as many parameters as possible.

  This criterion is designed to allow more definitive and accurate distinction of sources of contamination, to better explain differences in waste compositions, and to better explain relationships of waste composition. This is an overall PCA and investigative objective.

- Exclude redundant parameters.

  Parameters that measure similar attributes or composition of the samples were excluded from the PCA in most cases to avoid placing to much weight on similar constitutents. For example, conductivity was excluded in the water PCA runs because it measures the same attribute as total dissolved solids (TDS). In addition, dissolved metals were typically excluded in the water PCA runs in lieu of total metals since dissolved metals measure the same attribute and are typically a substantial portion of total metals. Metals (e.g., copper) tend to form complexes with the large amount of organic matter in the poultry waste (see Moore et al. 1998). Hence total metals, which include both complexed and colloidal forms, better represent the metal transport during field runoff and subsequent transport in streams. Use of total metals also avoids any problems associated with the small amount of samples where dissolved concentrations were reported higher than total concentrations (see section 3.10 for discussion). Sensitivity runs (see Step 14) were performed with both dissolved and total metals (either total or dissolved).

  Various forms of phosphorus were also excluded due to potential redundancy (and other reasons) in both the water and solids PCA runs. In the water runs, only three forms of phosphorus were retained: total dissolved phosphorus (filtered; P_TD), soluble reactive phosphorus (filtered; P_Sol_Reac), and total phosphorus (not filtered; P_T). These three forms of phosphorus were retained because they are the most important forms used in modeling and other evaluations, and because, though somewhat redundant, they may aid in distinguishing sources. In addition, selected phosphorus analytical methods were eliminated based on protocols established and documented in the main water and solids databases. In all cases, phosphorus by method 6010 was eliminated because it was shown to have interferences and resulted in inaccurate data (see Section 3.8). Even though phosphorus by method 6020 provided reliable results (see Section 3.8), it was redundant with total phosphorus (not filtered) and dissolved total phosphorus (filtered). In addition, phosphorus results by method 4500 (Standard Methods) were typically used in lieu of phosphorus results by method 365.2 because the detection limits were lower. See Section 3.8 for a more complete discussion and comparison of phosphorus methods. In addition, sensitivity runs were performed with only one of the phosphorus parameters (vs. three). This and other sensitivity run results are discussed further in Step 14.

- Exclude parameter results by unreliable methods.

  As previously discussed, phosphorus by method 6010 was eliminated because results were not accurate.

- Exclude parameters that were not routinely analyzed.

- Variables with low relative numbers of observations (counts) were not retained for the PCA. The basis for this criterion was to minimize the impact of missing data on

PCA, which affects the ability of the PCA to generate reliable PC scores. Retaining these parameters would create "holes" in the correlation matrix and statistical evaluations could not be performed. **Tables 6.11-4a** (water) and **6.11-4b** (solids) provides a list of parameters not routinely analyzed that were excluded from the PCA.

- Exclude parameters with a substantial amount of nondetects.

  Variables with relatively high percentages of nondetects (as indicated on the pplots or by calculations) were not retained also for the PCA. The basis for this criterion was two-fold: either such variables were considered of insufficient variance (i.e., constants) or they were deemed to have too few observations above analytical detection limits to be reliably used for the PCA. These variables were identified iteratively during previous, preliminary PCA runs, and hence were removed at the main database stage during the current analyses. **Tables 6.11-5** (water) and **6.11-6** (solids) provide the frequency of detection for each of the measured parameters that were retained and that were excluded. As shown for the water samples, the frequency of detection of all retained parameters was typically larger than 55 to 60 percent except for total arsenic (46% detections in water). Arsenic was retained for the water PCA runs because it is an important parameter in distinguishing poultry waste from other wastes (it is added to poultry feed). A sensitivity analysis was performed with and without arsenic (see Step 14). No significant differences were observed in the results. In addition, some of the dissolved metals (aluminum, iron and arsenic) have lower frequency of detections. For major runs, only total metal concentrations were used. In addition, sensitivity runs were performed using dissolved metals instead of total metals (see Step 14). For solids, the frequency of detection for retained parameters was typically above 70 percent expect for sodium, beryllium and staphylococcus.

- Select parameters with good variability and good distribution.

  Variables with low relative variability as indicated by their limited range (maximum - minimum) and/or small standard deviation were not retained for purposes of the PCA. The basis for this criterion was to minimize the impact on the PCA of variables with low or insufficient variance, since such variables were either not useful for the PCA or are considered constants (not variables). During Step 7, descriptive statistics (minimum, maximum, etc) and probability plots were evaluated. Probability plots (of transformed data as applicable) were examined visually to ensure that the measured concentrations had a good distribution (near linear plot with good variation of concentrations from low to high). Example probability plots are provided in **Appendix E**.

- Exclude parameters for which concentrations in the waste source are similar to background concentrations and as a result may not provide good variation in the environmental samples.

  For example, nickel in both poultry waste and background soils have similar concentrations. Originally (in previous, preliminary PCA runs), nickel was excluded from the PCA. However, based on the frequency of detection (60%), it was decided to retain nickel in subsequent analyses. Sensitivity analyses was performed during previous PCA runs with and without nickel to determine if any large differences were observed (see Step 14). No significant differences were observed in the results. All PCA runs for this report included nickel.

Based on the above criteria and evaluations, a maximum of 26 water parameters, and a maximum of 32 solids parameters, were selected for the various PCA runs. For some of the sensitivity and investigative runs (see Section Step 14), these numbers were lower (e.g., 24 parameters were selected in the water sensitivity runs using only one of the three phosphorus parameters). For the two main water PCA runs presented in detail in this report (SW3 and SW17), the parameters retained and included in the PCA were as follows:

Total Aluminum	Alkalinity
Total Arsenic	Total Barium
Total Calcium	Chloride
Total Coliforms	Total Copper
E. coli	Enterococcus
Total Iron	Fecal Coliforms
Total Potassium	Total Magnesium
Total Manganese	Total Sodium
Total Nickel	Nitrite + Nitrate
Soluble Reactive Phosphorus	Total Dissolved Phosphorus
Total Phosphorus	Sulfate
Total Dissolved Solids	Total Kjeldahl Nitrogen
Total Organic Carbon	Total Zinc

For one of the two main solids PCA runs presented in detail in this report (SD1), the parameters retained and included in the PCA were as follows:

Total Aluminum	Total Arsenic
Total Barium	Total Beryllium
Total Calcium	Total Cobalt
Total Coliforms	Total Chromium
Total Copper	E. coli
Enterococcus	Total Iron
Fecal Coliforms	Total Mercury
Total Potassium	Total Magnesium
Total Manganese	Total Sodium
Water Soluble Ammonium	Total Nickel
Total Nitrogen	Organic Matter
Phosphorus (Mehlich 3)	Total Phosphorus
Water Soluble Phosphorus	Total Lead

pH (1:1)	Soluble Salts
Water Soluble Sulfate	Staphylococcus
Total Vanadium	Total Zinc

For the second of the two main solids PCA runs presented in detail in this report (SD6), which included core samples collected from Lake Tenkiller, the parameters retained and included in the PCA were as follows:

Total Aluminum	Total Arsenic
Total Barium	Total Beryllium
Total Calcium	Total Cobalt
Total Chromium	Total Copper
Total Iron	Total Mercury
Total Potassium	Total Magnesium
Total Manganese	Total Sodium
Total Nickel	Total Nitrogen
Organic Matter	Total Phosphorus
Total Lead	pH (1:1)
Soluble Salts	Total Vanadium
Total Zinc	

The above list for the included core samples differs from the previous list (without the core samples) because the core samples were not analyzed for as many parameters. For example, the core samples were not analyzed for bacteria. Hence the PCA runs that included the core samples were reduced to a smaller number of variables.

Step 9: Normalize and Standardize Data and Perform PCA

As discussed in the previous Steps 7 and 8, typically all data used in the PCA were first normalized by using a log (base 10) transformation. In addition, standardization in the form of an autoscale (or z) transformation is conducted automatically by SYSTAT during a PCA run by analyzing a correlation matrix. The autoscale transformation, which ensures homogeneity of variance in the PCA, is defined as follows:

$$z_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i}$$

where z_{ij} is the datum (typically though not always base-10 log transformed) for variable i and sample j, and $\bar{x}_i$ and s_i are the mean and standard deviation, respectively, of the data (again, typically base-10 log transformed) for variable i and all samples.

As previously discussed, the EXCEL add-in program EDAnalyzer is used to facilitate the PCA. The EDAnalyzer program performs three primary functions: (1) interactive selection of groups, variables, and samples via distributional and data exploration analyses, including graphical displays; (2) generation of SYSTAT datasets and command files, and the running of the SYSTAT program via a shell application; and (3) managing and loading of SYSTAT result files for interactive graphical displays, along with options for saving selections and results files. These functions are described in further detail below. Note that EDAnalyzer is a utility program: all of its functions can be performed by hand prior to and following PCA analysis in SYSTAT; however, hand-performance of these functions is tedious and subject to mistakes, and therefore EDAnalyzer was used to ensure that a standard process was used.

Sub-databases (e.g., Subdatabase_Water_0428.xls) are loaded into EDAnalyzer. Using the fields: EDA_Group, EDA_Sample, and EDA_Variable, the user selects a set of records pertinent to the desired PCA. The selections are stored on the worksheet: Selections, which can be saved to a file in order to document the selections. Distributional and EDA is conducted via generation of a cross-tabulation, which is the rectangular (samples in rows and variables in columns) dataset that SYSTAT uses for actual PCA analysis (provided on the worksheet: Crosstab). The cross-tabulation is interactively examined to identify samples with a sufficient number of variables for PCA analysis. EDAnalyzer has an option for selecting a minimum number of variables, e.g., at least 20 of 26 variables. EDAnalyzer also has an option for creating (or averaging) the cross-tabulation by sample or by location; e.g., in the case of by location, the data for a particular variable with multiple samples assigned to that location would be averaged during creation of the cross-tabulation. For the current PCA runs in this report, no averaging is performed except when actual field splits (duplicates) were sampled, and no averaging is performed for samples collected at the same locations but at different times. In addition, EDAnalyzer has an option for selecting a multiplier for use on nondetect data: for all current PCA runs this multiplier was set to 0.5, meaning that values below analytical detection limits were set to one-half the detection limit.

As previously discussed in Step 7, distributional and data exploration analyses conducted in EDAnalyzer includes generation of a Pearson r correlation matrix (provided on the worksheet: Matrix). The correlation matrix is examined to ensure that (1) there will be no holes in the matrix, i.e., cases where a correlation cannot be calculated due to insufficient data and (2) there will be no cases with a correlation of 1, since that would indicate a condition where a variable was actually a constant in the PCA. Examination of the correlation matrix within EDAnalyzer is only a convenience in that SYSTAT can not perform PCA if the above conditions are not met – otherwise the correlation matrix generated in EDAnalyzer is not used directly by SYSTAT. As previously discussed, distributional and data exploration analyses conducted in EDAnalyzer also includes generation of various descriptive statistics and graphical displays (provided on the worksheet: Statistics). These are interactively examined and explored. The descriptive statistics are provided for each variable and include: Count, Minimum, Maximum, Mean, Median, Variance, and Standard Deviation. In addition, a normal probability plot is provided in order to allow examination of the distributional shape of the data and to assess the number of nondetects. EDAnalyzer has an interactive tool that allows the user to select various possible data transformation functions, including logarithmic, square, and square root, to view the transformed data on the probability plot, and then to save a selected transformation for subsequent inclusion as a command in the SYSTAT command file created by EDAnalyzer. The protocol used was to select the transformation that most closely "normalizes" the variables - typically this was the logarithmic (base 10) transformation. Finally, the probability plots and accompanying descriptive statistics are examined to ensure that the variable has sufficient variance (variability) to be considered a useful variable in the PCA. Variables with insufficient variance, or a large percentage of nondetects, are not useful variables and may cause the SYSTAT PCA to not execute.

Once interactive selection and data exploration are completed, EDAnalyzer creates the dataset (essentially a copy of the Crosstab worksheet) and the command file for the SYSTAT PCA. EDAnalyzer then shells out to SYSTAT via execution of the command file. The command file contains instructions to SYSTAT for creating and managing input and output files and for transforming variables, along with the detailed instructions for the PCA. The SYSTAT output files are stored within a fixed-location folder and they are always given the same names. The user ensures that the PCA run was successfully completed (all commands were executed) - if not, SYSTAT provides an error message. Following successful execution, the output files in the fixed-location folder (from a previous PCA run, if any) are overwritten.

Following completion of a successful PCA run in SYSTAT, control goes back to EDAnalyzer and the user then loads the SYSTAT results directly into EDAnalyzer into various worksheets: Standard - which contains the standardized data generated in SYSTAT and actually used in the PCA, and Results - which contains loadings, coefficients, percent variance explained, and PC scores for the first five principal components and for five different rotations. The PC scores are generated within EDAnalyzer using the coefficients and standardized data. Additionally, a re-scaled PC score is calculated for each sample or location. The Result worksheet also contains the EDA_Groups selected for the analysis.

Although EDAnalyzer only extracts results for the first (or top) five principal components, SYSTAT actually generates results for all possible principal components, one for each variable. EDAnalyzer only shows the results for the first five principal components because it is rare that information in components beyond the first 2-3 is useful in environmental studies.

After the Results worksheet is populated, EDAnalyzer uses this information to generate various PCA graphical displays (provided on the worksheet: Display). The graphical displays include: horizontal bar charts showing both the loadings and coefficients of the parameter for the first two PCs, a vertical bar graph which shows the percent variance explained by each of the five principal components, a PC by PC scatter plot showing the loadings with variable labels, a PC by PC scatter plot showing PC scores with sample/location labels, and a map (X versus Y coordinate) showing the sample/location points size-scaled according to the value of the PC score selected. EDAnalyzer provides an interactive tool that allows for selecting various principal components and rotations for graphical display.

The numerical PCA results on the Results worksheet can be and typically are saved to separate files in order to document the PCA and to save the results for subsequent analysis, graphical display, and mapping purposes. The graphical displays of the results generated within EDAnalyzer are not typically saved; however, EDAnalyzer can import saved results files in order to display them graphically at a later time. For all current PCA runs presented in this report, both the cross-tabulated dataset and the PCA results were saved to individual files. These files were given names corresponding to the date the PCA was conducted, e.g., Crosstab_Water_0427_SW_3.xls and Results_Water_0427_SW_3.xls are the saved cross-tabulation and results files for water PCA run SW 3 conducted on April 27, 2008. The sub-database loaded into EDAnalyzer, used in making a PCA run, and pertaining to the cross-tabulation and result files, is the file with the same or most recent previous "date" (e.g., for the example, this was the sub-database file named Subdatabase_Water_0427.xls).

The cross-tabulation and result files generated by EDAnalyzer and saved (as described above) were subsequently used to generate various additional files for data analysis and graphical display, depending on the current needs and level of analysis (e.g., sensitivity or investigative analysis). Generally, these additional files included a file with prefix "R_PC_Plot" that contains the PC 1 through PC 5 scores for all five rotations, along with a series of PC by PC scatter plots pertaining to all five rotations. These and other files were also generated in order to provide graphical displays and tabulations presented in this report.

Step 10: Identify Major Principal Components

The total variability (or variance) in a multivariate dataset is a function of the number of parameters and their individual variances. If the parameters exhibit no inter-relationships or mutual correlations then the proportion or percentage of the total variance explained by or accounted for by each variable (parameter) would be the same. For example, the percentage of the total variance accounted for by each variable in a dataset with $i = 26$ parameters, given no mutual correlations, would be $(1/i) \times 100 = (1/26) \times 100 = 3.85\%$. However, this is not true for a multivariate dataset where the parameters exhibit at least some degree of mutual correlation. Principal components analysis (PCA) is a commonly used multivariate statistical method for identifying these mutual correlations, if present, and re-apportioning the individual variances accordingly.

PCA operates by transforming a dataset with a large number of parameters, ostensibly with inherent mutual correlations, to a new set of uncorrelated reference parameters called principal components or PCs. The number of PCs is the same as the number of original parameters. However, the apportionment of the total variance among the PCs will depend not only on the number of PCs but on the mutual correlations exhibited by the original parameters that comprise the PCs. Given mutual correlations, the objectives of PCA are to: (1) identify those PCs that explain or account for relatively high percentages of the total variance in a dataset, and (2) examine these PCs in order to interpret meaningful relationships among the samples in the dataset. These objectives can only be met by PCA in those cases where the parameters exhibit mutual correlations—and hence the dimensionality of the parameters in a multivariate dataset can be reduced to a smaller number of significant PCs—and where these PCs exhibit relationships among the samples from which meaningful interpretations are possible. The term "significant" in this context means that a relatively high percentage of the total variance is accounted for (explained) by a small number of PCs.

Experience has shown that the objectives of PCA can be met in a dataset or environmental system dominated by a relatively few number of source impacts that exhibit mutual correlations among their parameters. In such cases a correspondingly high percentage of the total variance is explained by only a few PCs, typically 2-3 PCs. This is the reason why EDAnalyzer only extracts (for examination) the top or most significant five PCs: if the top five PCs do not account for a high percentage of the total variance in the system then there is little hope of interpreting meaningful relationships among the samples. In PCA, the PCs are sorted according to the percentage of total variance explained, i.e., from those PCs that account for the highest percentage to those that account for the lowest percentage. One then examines these percentages in order to identify the significant PCs, if any.

Many different PCA runs were conducted during this investigation, some of which have been classified as "sensitivity" analyses and some of which have been classified as "investigative" analyses. Those classified as sensitivity analyses were designed to evaluate the sensitivity on the PCA of using certain different parameter sets or sample groups. The sensitivity analyses and their results are discussed in more detail in Step 14. The investigative analyses were designed for more direct analysis and interpretation relative to identification of source signatures in the watershed. From the investigative PCA runs, four have been selected (two for water samples labeled SW 3 and SW 17, and two for solids samples labeled SD 1 and SD 6) as the most important to the investigation or project objectives. Hence the results of these four PCA runs are presented in detail in this report. Aside from their importance, these four runs are also representative of the method used to examine the significance of the PCs, as discussed above, and therefore will be used as such in this section.

One method of displaying the significance of the PCs graphically is a point plot of the percent variance explained versus each PC, where the number of PCs is equal to the total number of original variables—and hence one can show how the variances differ from a corresponding alternate case of no mutual correlations. Such a plot is known as a scree plot, the term "scree" meaning "rubble at the bottom of a cliff" and referring to the random noise in the dataset as the number of PCs increases. In this context, "random noise" refers to the variance attributable to the original variables (parameters) and unrelated to the significant PCs. **Figure 6.11-1** shows a scree plot for PCA run SW 3, which contained 26 variables and hence corresponds to 26 PCs, PC 1 through PC 26 on the plot. As shown, the top five PCs (PC 1 through PC 5; indicated with blue symbols) each account for more than $(1/i) \times 100 = (1/26) \times 100 = 3.85\%$ of the total variance in the dataset, the amount attributable to random noise or to each original variable, and hence are considered significant. The amount of variance actually explained by the top five PCs for SW 3 is 74.1%, a significant proportion of the total variance and a significant reduction in dimensionality: from 26 variables explaining 100% of the variance to 5 PCs explaining 74.1%. The remaining variance, (100 - 74.1) × 100 = 25.9%, is considered to be random noise and is unrelated to the first five PCs. An alternate way of displaying this same information is a scree plot in the form of a bar graph, as shown in **Figure 6.11-2** for SW 3. On the bar graph, the percentage of the total variance explained by the top five PCs are each indicated, i.e., 38.0% (PC 1), 18.2% (PC 2), 7.6% (PC 3), 5.3% (PC 4), and 4.9% (PC5). These variances indicate that PC 1 and PC 2 are by far the most important of the five together explaining 56.2% of the total variance, relative to PCs 3, 4, and 5 (17.8%). Similar plots for the other PCA runs are provided in **Figures 6.11-3** through **6.11-8**. These all clearly show that the top five PCs are significant (above random noise), and that the top two PCs are the most significant. These results were used to establish the top two PCs (PC1 and PC2) as representing the dominant signals or signatures related to impacts in the watershed. The dominant PC1 and PC2 signatures also proved to be interpretable as to source identification because they are so dominant - see steps 12 and 13. On the other hand, PCs 3, 4, and 5 generally were less readily interpretable (because they are so much closer to random noise or background variation).

For the two water PCA runs (SW 3 and SW 17), there is no particular advantage of one scree plot version over the other: they both show the same information. However, for the two solids PCA runs (SD 1 and SD 6), the bar graph version has the advantage of also showing an alternate PCA rotation (called the varimax rotation) that proved useful for additional interpretation of the sample PC scores, as discussed further in this report. The objective of varimax rotation is to use the significant PCs (in this case PC 1 through PC 5, i.e., ignoring the insignificant PCs 6 through 26) and rotate or adjust their PC axes to maximize the variance of the variable loadings (closer to +1 or -1). This rotation proved to assist the interpretation in terms of the original variables and to allow more definitive distinctions of PC scores, in the cases of the solids PCAs. As shown on the corresponding figures, the varimax rotation apportions the percentage of total variance differently; however, the total variance explained by the top five PCs is the same: in the case of SD 1, 81.4%, and in the case of SD 6, 81.7%. Again, as in all PCA runs, the PCA in all cases successfully reduced the dimensionality of the datasets from a large number of original variables to a relatively few significant PCs, hence allowing for meaningful interpretations of source impacts in the watershed.

A summary of the variance explained by PC1 and PC2 for each of the four major PCA runs are shown below:

Run	Groups	Rotation	Variance Explained by PC1 (%)	Variance Explained by PC2 (%)	Variance Explained by PC1 and PC2 (%)
SW 3	Surface Waters	No Rotation	38.0	18.2	56.2
SW 17	Surface Waters and Groundwaters	No Rotation	34.2	15.9	50.1
SD 1	Solids (wastes, soils, sediments)	Varimax	38.3	16.7	55.0
SD 6	Solids and Core Samples	No Rotation	38.5	28.5	67.0

Each calculated PC (significant or otherwise) has associated with it a set of coefficients that relate the value of the PC to the values of the original variables. Hence these coefficients can be multiplied by the values of the original variables, and then summed, to calculate a PC score for each sample in the dataset. The method of calculating PC scores, and how these scores are used in evaluating the samples, is discussed in further detail in Step 11. The variance of the values (PC scores) of a particular PC for all samples in a dataset determines what is equal to a quantity called an eigenvalue, which is equal to the variance of the PC and therefore used to calculate the percent variance explained by the PC. For example, for PCA run SW 3, which contains 26 variables, the eigenvalue for PC 1 is 9.89, and therefore the percentage of the total variance explained by PC 1 is $(9.89/26) \times 100 = 38.0\%$, as is shown in **Figure 6.11-9.** This illustrates how these scree plots are generated.

The correlation between the values of the PCs for all samples and the corresponding values of the original variables is called a loading. A loading is a re-scaled coefficient such that they become correlation coefficients. Hence it is useful and meaningful to examine the loadings (or the coefficients) in order to determine the importance of the original variables for a particular PC. This is a key step in interpreting the PCs with regard to source signatures, as those variables with relatively high loadings (significant correlations) may be related in terms of geochemical mechanisms and transport pathways to similar high concentrations (or correlations) in the waste source. The actual interpretations of the PCs are presented and discussed in Step 12 of this report. The loadings and coefficients for four critical PCA runs (SW 3, SW 17, SD 1 and SD 6) are provided in **Figures 6.11-10** through **6.11-17**. **Figure 6.11-10** provides bar graphs of the loadings for PC 1 and PC 2 for PCA run SW 3. As shown, PC 1 exhibits relatively high positive loadings (greater than 0.6) for a large number of variables, including: arsenic, total coliforms, copper, e. coli, enterococcus, iron, fecal coliforms, potassium, nickel, total and total dissolved phosphorus, total organic carbon, and zinc. These are interpreted as the most important variables comprising

PC 1, and therefore if these or a subset of these variables can be shown to be related to a particular waste source, then samples with high PC 1 scores can be related to, or have signatures consistent with, that source. Similarly, PC 2 exhibits relatively high loadings for a different set of variables: chloride, sodium, and sulfate, which may indicate a relationship to another source. The threshold loading (0.6) in this example is arbitrary and has been selected solely for illustrative purposes: such thresholds are commonly adjusted based on additional information available to the investigator. **Figure 6.11-11** provides the bar charts for Run SW 3 for both PC1 and PC2 with the coefficients shown instead of the loadings. **Figures 6.11-12** and **6.11-13** provide the loadings and coefficients for PC1 and PC2 for SW17. **Figures 6.11-14a** and **6.11-15a** provide the loadings and coefficients for SD1. **Figures 6.11-14b** and **6.11-15b** provide the loadings and coefficients for SD1 using the varimax rotation. **Figures 6.11-16** and **6.11-17** provide the loadings and coefficients for PC1 and PC2 for SD6.

Step 11: Calculate Principal Component Scores

Principal component (PC) scores are calculated for each identified significant PC for each individual sample. Identification of significant PCs was discussed in Step 10. To calculate a PC score for each individual sample, the PC coefficient is multiplied by the standardized parameter concentration. This is performed for all parameters (variables) in a particular PCA run. The product values for all 25 parameters are summed to yield one PC score for each sample for each PC. Hence, a particular sample will have both a PC 1 score and a PC 2 score. If one of the variables selected in a particular PCA run is missing a value (due to it not being measured), the product (coefficient times the standardized concentration) for that parameter is essentially not used in the summation: this is the same as multiplying the coefficient by the standardized mean concentration which is zero. Sensitivity runs were performed using datasets with no missing value (Step 14)

Once the PC scores have been calculated for each significant PC, they are examined graphically via PC-by-PC scatter plots. Since EDAnalyzer extracts (for examination) the top five PCs, the number of scatter plots produced for possible examination will be: (5)(4)/2 = 10, i.e., PC 1 vs. PC 2, PC 1 vs. PC 3, PC 1 vs. PC 4, PC 1 vs. PC 5, PC 2 vs. PC 3, PC 2 vs. PC 4, PC 2 vs. PC 5, PC 3 vs. PC 4, PC 3 vs. PC 5, and PC 4 vs. PC 5. Furthermore, since the PCA is conducted using five different possible rotation variations: no rotation, varimax, equimax, quartimax, and oblimin, a total of: (10)(5) = 50 PC scatter plots were actually produced.

PC1 vs PC2 plots are provided for the following PCA runs:

- SW3 (Surface Water)
  - Figure 6.11-18a
  - Figure 6.11-18b (expanded view)
  - Figure 6.11-18c (shows two major groups – WWTP impacted waters and poultry waste impacted waters)
  - Figure 6.11-18d and e (sample types identified)

- SW17 (Surface Water and Groundwater)
  - Figure 6.11-19a
  - Figure 6.11-19b (expanded view)
  - Figure 6.11-19c and d (sample types identified)

- SD1 (Solid Samples)
  - Figure 6.11-20a
  - Figure 6.11-20b (expanded view)
  - Figure 6.11-20 c and d (sample types identified)
  - Figure 6.11-20e (PC2 vs PC3)
  - Figure 6.11-20f (PC1 vs PC2, no rotation)

- SD6 (Solid Samples including Lake Tenkiller core samples)
  - Figure 6.11-21a
  - Figure 6.11-21b (expanded view)
  - Figure 6.11-21c and d (sample types identified)

Examination of the PC scatter plots is a key step with regard to interpreting source signatures in the watershed: the analyst seeks to identify patterns, groupings, and relationships in the PC scores that distinguish the samples based on the various waste source impacts. This is discussed in further detail in Steps 12. The PC scores were also mapped in order to examine spatial and temporal relationships of the samples to the various waste sources. The PC scores typically range from negative to positive values. In this investigation, mapping was facilitated by re-scaling the PC scores such that the lowest score for a particular PC was assigned a value of one. This is also discussed in further detail in Steps 12.

Principal Component Scores are provided in **Appendix F** for all four major PCA runs conducted during this investigation and discussed in this report. There were a total of 22 water PCA runs (SW 1 through SW 22) and eight solids PCA runs (SD 1 through SD 6). **Tables 6.11-7a (Water) and 6.11-7b (Solids)** provide a summary of the PCA Runs.

Step 12: Evaluate Whether Major Components are Associated with Specific Sources

This step consists of two evaluations: 1) comparison of the principal component parameters to the composition of known waste sources and 2) a spatial and temporary analysis of individual principal component scores (for all major principal components).

Comparison to Known Waste Sources
In section 6.4.2, the chemical composition of cattle manure, poultry waste and waste water treatment plant discharges were provided (taken from literature values). **Tables 6.11-8, 6.11-9** and **6.11-10** provides the compositions of the PCA parameters for the following materials collected from the IRW:

- For Solid Samples (**Table 6.11-8**): 32 parameters
  - Average composition of 16 poultry waste samples
  - Average composition of 5 fresh cattle manure samples and 5 dry cattle manure samples
- For Synthetic Precipitation Leachate (SPLP) Samples (**Table 6.11-9):** 25 parameters
  - Average composition of two poultry waste leachates
  - Average composition of five fresh and five dry cattle manure leachate
  - Note, because the SPLP procedures require filtering, no total P was reported and all metals are dissolved concentrations (25 parameters versus 26).
- For Liquid Samples (**Table 6.11-10**): 26 parameters
  - Average composition of runoff from fields with poultry waste application (60 samples) - note, fields also had some cattle manure
  - Average composition of runoff from fields with potentially only cattle manure (two samples)
  - Average composition of two springs documented with cattle manure
  - Average composition of WWTP discharge from samples collected at Rogers discharge, Siloam Springs discharge and Springdale discharge (note - all samples were collected during high flow rates because of infiltration to lines after rain)
  - Average composition of surface water samples (25 samples) impacted by and collected downgradient of WWTP discharges

As shown in **Table 6.11-8**, the parameters highlighted in yellow have a different composition when compared to poultry waste. Parameters where poultry waste and cattle manure have distinctly different concentrations (by a factor of at least 3 times) are copper, phosphorus, potassium, zinc, manganese, arsenic, sulfate, sodium, calcium, soluble salts, nickel, aluminum, chromium, and some bacteria. **Figure 6.11-14b** provides the PC1 parameters and loadings sorted in order of importance for the solid samples (run SD1) including poultry waste, cattle manure, soil (0-2 inch), river sediments and Lake Tenkiller sediments (grab samples only). The parameters with the largest loadings and most importance in the PC (shown by the length of the vertical bars) are at the top of the figure while the loadings with the lowest coefficients and least importance are at the lower part of the figure. As shown, 24 of the 32 parameters have positive coefficients. Sixteen (16) of the parameters have significant loadings above 0.5. Parameters with the largest loadings in PC1 in order of importance follow: potassium, total P, sodium, magnesium, water soluble sulfate, total zinc, soluble salts, Mehlich 3 P, copper, calcium, organic matter, water soluble ammonia, water soluble P, total nitrogen, enterococcus and e. coli. As shown in **Table 6.11-8**, many of these parameters have very large concentrations in poultry waste and relative lower concentrations in cattle manure including potassium, phosphorus, sodium and sulfate.

Most important, the PC1 score vs PC2 score figure (**Figure 6.11-20a** and **c)** shows that the cattle manure plots on the figure in a distinctly different group than the poultry waste. These two groups are most clearly separated using the varimax rotation. However, the separate groups are also observed on the PC1 vs PC2 figure using no rotation (**Figure 6.11-20f**). These figures show that cattle manure and poultry waste have different and distinct chemical/bacterial signatures.

Table 6.11-9 compares the synthetic precipitation leachates from poultry waste and cattle manure. The parameter concentrations highlighted in yellow have distinctly higher concentrations in poultry waste leachate than cattle manure leachate by a factor of at least 3 times. These parameters include: copper, iron, TOC, nickel, potassium, zinc, arsenic, total dissolved P, soluble reactive P, TKN, total dissolved solids, sulfate, chloride, sodium and alkalinity. **Figure 6.11-10** provides the PC1 parameters and loadings sorted in order of importance for surface water samples. As shown, 22 of the 26 parameters have positive loadings. Nineteen (19) of the 26 parameters have loadings above 0.5. Parameters with the largest loadings in order of importance include: copper, e. coli, iron, TOC, total P, aluminum, nickel, fecal coliform, enterococcus, total coliform, potassium, zinc, manganese, arenic, total dissolved P and soluble reactive P. As shown, poultry leachate has very high concentrations of all of these parameters. A PAC run was performed with both poultry waste SPLP leachate and cattle manure SPLP leachate (SW18 - see **Appendix F**). This run shows that the poultry waste SPLP and the cattle manure SPLP samples are in distinct groups. No runs were performed with the SPLP poultry waste samples and surface water samples because the very high PC scores for the SPLP sample would dominate the analysis.

Table 6.11-10 provides the concentration information for liquid (water) related wastes including edge of field, WWTP discharges and surface waters impacted by WWTP discharges. As shown by concentrations highlighted in yellow, the chemical composition of runoff from poultry waste applied fields is different than runoff from fields with only cattle manure. All parameters with measured concentrations are different by a factor of 3 or more. **Table 6.11-10** also provides the chemical composition of WWTP effluent samples and for samples collected in streams (25 samples) directly downgradient of WWTP discharges. As shown, the chemical and bacterial composition of runoff from poultry waste applied fields is distinctly different when compared to the WWTP effluent or stream samples. Different (much higher) chemical and bacteria concentrations include: cooper, e. coli, iron, TOC, total P, aluminum, nickel, fecal coliform, enterococcus, total coliform, potsssium, zinc, manganese, arsenic, total dissolved P, soluble reactive P, TKN, and barium. These parameters have very high concentrations in runoff from fields with poultry waste and leachate from poultry waste. **Table 6.11-10** also show that springs (two samples) impacted with cattle manure have a different composition and lower concentrations than runoff from fields with poultry waste or poultry waste leachate for most parameters including copper, e. coli, iron, TOC, aluminum, nickel, fecal coliform, enterococcus, total coliform, zinc, manganese, arsenic, TKN and nitrite + nitrate.

Figure 6.11-10 shows the loadings for the 26 parameters for both PC1 and PC2 for surface water samples (SW3). As shown for PC1, 22 of the 26 parameters have positive loadings and 19 of the parameters have loadings greater than 0.5. All of these parameters have very large concentrations in runoff from fields with poultry waste and leachate from poultry waste. **Figure 6.11-10** also shows the loadings for the 26 parameters for PC2. As shown, 14 parameters have a positive loadings and 7 have loadings larger than 0.5. The largest loadings in order of importance follow: sodium, chloride, sulfate, soluble reactive phosphorus, calcium, total dissolved phosphorus, potassium, magnesium, alkalinity, TDS and nitrite+nitrate. Of these parameters, calcium, sodium, chloride, nitrite+nitrate, and sulfate have larger concentrations in WWTP associated samples then in samples associated with poultry waste.

Because of the chemical and bacterial comparison discussed above, PC1 has been identified as associated with poultry waste and PC2 has been identified as associated with WWTP effluent. These identification were be confirmed by the spatial analysis discussed in the next section.

Spatial Analysis

The spatial/temporal analysis evaluated principal component scores in relation to the location of the sample (distance from sources), group type or environmental component (e.g, edge of field), sample conditions (e.g., high flow, base flow), poultry house density, and reference locations.

Appendix F provides the PC1 scores for the surface water samples (SW3) sorted from high to low values. The following observations can be made:

- The highest PC1 scores are the edge of field samples collected as runoff from fields with poultry waste application. Of the top 50 samples with highest PC1 scores (scores above a value of 2), 44 are edge of field samples. Four other samples in this group were collected at USGS stations or small tributarties stations during very high flow conditions. The highest PC1 score is 8.1 for an edge of field sample collected after documented poultry waste application and from water flowing off the field. The fact that the highest PC1 scores are from the edge of field samples is consistent with the samples being collected at the source of surface water contamination; i.e., the runoff from fields with poultry waste. These are the locations where the most PC1 parameters were detected at the highest concentrations.

- The lowest scores are from reference areas or areas with minimal poultry houses and operations. The lowest score (1.00) is from REF2 (Dry Creek), the only true reference with no poultry houses in the area. Other "reference" locations outside the IRW, REF1 and REF3 (Little Lee Creek and Spring Creek) have the third and fifth lowest PC1 scores, respectively (1.18 and 1.19). Other low scores were from samples collected at HFS30 and HFS28A which are small tributarties in the IRW with low poultry house density. Some low scores were also observed for some USGS stations on the Baron Fork and HFS20. HFS20 has low poultry house density in the actual basin, but high poultry house density within a two mile radius. If PC1 represents poultry contaminant, then areas with minimal poultry impacts should have the lowest PC1 scores.

- **Figures 6.11-22a and 6.11-22b** show box plots with the median, lower quartile and upper quartile for the PC1 scores for the following groups: edge of field samples, small tributartiy locations with samples collected at high flow, small tributarty locations with samples collected at base flow, USGS stations (at high flow), USGS stations (base flow), surface water samples collected at biological and other river locations (mostly base flow), samples collected in Lake Tenkiller and samples collected at reference or locations with minimal poultry waste impact. As shown the median and upper quartile PC1 scores typically decrease in value in a logical order according to the known pathways from very high at the edge of field to very low at the reference locations. After edge of field samples, samples collected during high flow conditions in the small tributarties have the next highest scores followed by base flow samples collected at the same locations and surface water samples collected at high flow conditions. The median PC1 score for USGS samples collected at high flow show an increase compared to the median for surface water samples collected for other river samples. The PC1 scores for samples collected from Lake Tenkiller are higher than the PC1 scores for samples collected at the USGS stations during base flow conditions. The reference areas have the lowest PC1 scores. This evaluation shows the transport of PC1 parameters from the edge of field to rivers and streams and finally to Lake Tenkiller.

Appendix F shows the PC2 scores sorted from the highest to lowest scores for run SW3. Several observations can be made:

- Of the highest 65 PC2 scores (above PC2 values of 4.8), three are discharge samples from WWTPs, 52 are surface water samples and 10 are the anomalous EOF samples discussed in Section 6.8. Of the 52 surface water samples, 48 are downgradient of WWTP discharges. This includes 18 samples at HFS04 (downgradient of Siloam Springs WWTP discharge) and 16 samples at HFS22 (downgradient of Lincoln WWTP discharge). Samples from locations 345, 121, 75, 349, 31, 350, 901, 120, 109, 72, 122 and 246 are also in this group. These samples are downgradient of discharges from Rogers, Springdale, Siloam Springs, Prairie Grove, Lincoln, Westville and Fayetteville WWTP discharges. Most of the samples are downgradient of Springdale or Rogers. See **Table 6.11-11** for the largest PC2 scores and locations.

- Of the highest 65 PC2 scores, 10 are from edge of field samples. However the chemical/bacterial compositions of these 10 samples are distinctly different than effluent from WWTPs and are discussed in detail in Section 6.8. These 10 samples also have very high PC1 while the WWTP impacted samples do not have high PC1 scores. These 10 samples are not WWTP effluent impacted but are thought to be fresh leachates collected during very high runoff conditions. These samples could potentially contain both cattle manure and poultry waste contamination.

Summary Observations

Because of the spatial analysis and comparisons to waste compositions, PC1 has been identified as related to poultry contamination (i.e., a poultry waste signature) and PC2 has been identified as related to WWTP discharge (i.e., a WWTP signature). In addition, high PC1 scores are observed along the major flow pathways and are higher near sources of poultry waste land application and decrease with distance from the source areas. The evaluation of these observations is performed in conjunction with the next two Steps of the PCA evaluation: step 13 (Use of PC Scores to Determine Sample and Locations Impacted by Major Sources of Contamination) and step 14 (Investigative and Sensitivity Runs).

Step 13: Use the PC Scores to Determine the Samples and Locations in the IRW that are Impacted by Major Sources of Contamination

As previously discussed in Step 12, a spatial evaluation was performed to evaluate the individual sample PC scores in relation to distance from sources, sample group, sample conditions and reference locations. In this step the individual PC scores were evaluated to determine the magnitude of impact or contamination from sources across the basin. If contamination is pervasive and dominant across the IRW in all environment components, a pattern or signature groups of each major source of contamination should be observed when evaluating PC scores relative to each other.

Figures 6.11-18a and 6.11-18b provides a plot of the PC1 (x-axis) vs the PC2 (y-axis) scores for run SW3. **Figure 6.11-f** shows all 573 scores and **Figure 6.11-18b** shows only the scores for the samples inside the box shown in **Figure 6.11-18a** ("Area of Expanded View"). **Figure 6.11-18c** shows all points in the expanded view area (560 out of the 573 samples are shown). The figure also shows lines around the two major groups of samples identified from PC1 and PC2 evaluations. The group with high PC1 scores is labeled "poultry dominant impact" and contains the samples whose chemical and bacterial composition is dominated by poultry contamination. The group with high PC2 scores is labeled "WWTP dominant impact". These are the samples in which the WWTP impact or influence on the sample is greater than the poultry impact. There are 57 samples in this group (10 % of total). It is important to note that except for some of the reference samples, most of the samples (even those "dominated" by WWTP) show some poultry contamination.

The two groups were selected by examining the locations and chemistry/bacterial composition of the individual samples. For the "WWTP dominant impact" group, the PC2 scores were selected to be above a value of 4.7. As shown in **Table 6.11-11**, samples below about a score of 4.8 are typically not in locations downgradient of WWTP discharges so cannot be impacted by WWTPs. For the "poultry waste dominant impact" group, a PC1 score of greater than 1.3 was selected. This is a conservatively high value and could have been set lower to include more samples. The value was selected by examining the locations and scores of samples, particularly the scores of reference samples and samples in low poultry density areas. In summary, the samples with PC1 scores below approximately 1.3 include all samples from reference locations (six total), 9 out of 10 samples from HFS30 (small watershed location with low poultry house density) and 10 out of 11 samples from HFS28A (small watershed location with low poultry house density). The one sample from HFS30 and the one sample from HFS28A with higher PC1 scores were collected during extreme flow events. Overall, 441 of the 573 samples (77%) had PC1 scores higher 1.3 and show some poultry contamination.

Figure 6.11-23 shows the average PC1 scores by location (based on PCA run SW3). The average PC1 score was determined if multiple samples were collected and contained in the PCA analyses by calculating the mean score of those samples. In **Figure 6.11-23**, there are 175 different locations. Of these, 137 have a PC1 average scores greater than 1.3. Therefore, approximately 78 percent of the locations sampled in the IRW show some poultry contamination. Locations with PC2 scores higher than 1.3 are shown in red; those with scores less than 1.3 are shown in green.

The following table gives a breakdown of the number of samples with poultry contamination by the various sample types (based on run SW3):

Sample Type	Sample Counts	Percent > 1.3
EOF	65/65	100
Lake Tenkiller	29/29	100
Steam – base flow	56/90	62
Stream -high flow	13/20	65
Small Trib-base flow	32/48	67
Small Trib-high flow	158/177	89
USGS – base flow	32/48	67
USGS – high flow	60/81	74

Note: the three WWTP discharges samples are not included because they are actual source samples; reference samples are included in the "streams" group.

Evaluation of Groundwater and Spring Samples

Figures 6.11-19a and 6.11-19b show the PC1 score vs PC2 score plot for PCA run SW17. This run is the same as SW3 except groundwater samples (geoprobe and existing wells) and springs samples are included in the PCA. This results in 699 total samples in the PCA. The results of this run are provided graphically and include:

- **Figures 6.11-3** and **6.11-4:** Scree Plots and Variance Analysis
- **Figures 6.11-12** and **6.11-13**: PC Parameters, Loadings and Coefficients

- **Figures 6.11-19a, b, c** and **d**: PC1 vs PC2 plots

In addition, **Figure 6.11-22c** provides box plots showing the PC1 scores for geoprobe samples, spring samples and existing well samples (run SW17). As shown, there is a decrease in the median PC1 values with Geoprobe samples having the highest PC1 scores, than springs and existing wells have the lowest PC1 scores. This is a logical progression from shallow alluvial water to springs and to deeper wells.

A similar evaluation of PC1 scores was performed for the SW17 run as for the SW3 run where the PC scores for reference samples and samples from locations in areas of low poultry house density were evaluated. This resulted in determination that the same threshold PC1 score could be used to determine poultry waste impact (samples with PC1 > 1.3). The locations of the springs, wells and geoprobes with PC1 average values above and below a value of 1.3 are shown in Figure 6.11-24 (based on PCA run SW17). There are 112 locations on the figure and 51 have PC1 values of greater than 1.3 (red dots). These locations are impacted with poultry contamination (46 percent). The following table shows the number of individual samples with poultry contamination (run SW17):

Sample Type	Sample Counts	Percent > 1.3
Geoprobe	16/17	94
Springs	19/49	39
Existing Wells	24/60	40

Overall, 59 out of 126 geoprobe, springs and well samples (47%) show poultry contamination. The three wells known to be greater than 150 ft in depth (actual depth = 203 to 803 ft) did not show poultry waste contamination. Four of the grower's wells (unknown depth) did show poultry waste contamination. Sample locations with PC1 scores reflecting poultry waste contamination are located through out the Oklahoma portion of the IRW (most all sample locations where in Oklahoma) and demonstrate that contamination is widespread for residential wells and alluvial groundwater.

In addition to the samples showing poultry waste impact, some of the groundwater samples have higher PC2 scores than the typical samples identified as being impacted with poultry waste contamination (relatively lower PC2 scores). These groundwater samples potentially show human waste impact. Overall about 20 wells may show potential human impact.

Evaluation of Potential Impact of Cattle Manure

The potential impact due to cattle manure was previously discussed in Section 6.4.2. These mass balance calculations indicate that any impact or contamination from cattle manure would be small (typically < 10 percent of the mass for most chemical constitutents) compared to the impact due to poultry waste disposal. Previous steps n this subsection (i.e., step 12 discussing waste characteristics) show that cattle manure and cattle manure leachate are very different in chemical composition when compared to poultry waste and poultry waste leachate. Therefore if cattle waste provides a major impact on contamination in the IRW, a dominant signature should be observed in the PCA. To assist in this evaluation, samples with known cattle contamination were evaluated. The chemical and bacterial compositions of these samples have been previously provided in **Tables 6.11-10 and 6.4-2a)**. The four samples documented with cattle contamination are: SPR-LAL16-SP2, SPR-26, EOF-CP-1B and EOF-CP-1A. **Figure 6.11-25** shows the PC1 vs PC2 score plot for PCA run SW22 (surface water and springs; same as SW3 with springs added). Also shown on this figure are the locations of the four samples with potential cattle contamination (red dots). One of the spring samples (SPR-26) plots in the WWTP impact area and another spring sample (SPR-LAL16-SP2) plots above the WWTP impact area. Field notebooks indicate that SPR-LAL16 was definitively contaminated with cattle manure while SPR had the potential for cattle contamination. The other two samples plots near the edge of the poultry waste impacted area. These four samples have very different PC scores and no consistent relation or group is observed in the PCA. If cattle contamination contributed a significant impact to contamination in the IRW, a clear signature and associated group should be observed in the PCA and the four samples with cattle contamination would be in the group. Based on the mass balance calculations, the comparison of chemical composition and the PCA analyses, cattle waste is not a major source of chemical contamination in the IRW.

Evaluation of Solid Samples

As previously discussed in Step 12 and shown in **Figure 6.11-20a**, cattle manure and poultry waste samples form two distinct groups (PCA run SD1, varimax rotation). In addition, soil samples (0-2 inches) collected from poultry waste applied fields and sediment samples are typically more closely related to poultry waste samples than to cattle manure samples. This shown in **Figure 6.11-20e** (run SD1, varimax, PC2 vs PC3) where the cattle waste is distinct from the soils and sediments samples. The poultry waste samples are closely related the soil and sediment samples.

Both PC1 and PC2 have high loading parameters that are related to poultry waste contamination. **Figures 6.11-20a, b, c** and **d** provide the PC1 vs PC2 plots of run SD6 (solid samples including Lake Tenkiller core samples, no rotation). **Figures 6.11-20b** and **6-11-20d** show an expanded view of the PC1 vs PC2 plots. The core samples typically show a decrease in PC2 scores from the shallow (more contaminated samples) to the deeper (less contaminated samples). As has been previously discussed (see section 6.7.2), this contamination in the Lake Tenkiller core samples is the result of poultry waste. As shown **Figures 6.11-20b** and **d**, these contaminated core sample plot with most of the soil and other sediment samples collected from the IRW.

Step 14: Perform Investigative and Sensitivity Analyses

Analyses were performed to evaluate the change in the PCA results due to various database selections or to determine the "sensitivity" of the results due to change in various elements of the PCA. In particular the change was evaluated by comparing the PCA results between various PCA runs. The results evaluated included comparison of the magnitude of the parameter coefficients, the percent variance explained and the PC scores for the individual samples. Changes made in the PCA runs included the number of parameters, specific parameters (e.g., arsenic and nickel), the groups or types of samples from environmental components (e.g., combinations of different environmental components), types of analyses (e.g., various forms and analytical methods for phosphorus) and specific samples (e.g., outliers).

In particular, the following sensitivity runs were previously performed:

- Surface water samples with and without additional phosphorus parameters. Retention of three form of phosphorus may be redundant and bias result to those samples with phosphorus. Similar runs were also performed for this current report.

- Surface water samples with and without the following parameters: arsenic, nickel, nitrate+nitrite and alkalinity. These were the parameters which were on the border line based on the parameter selection criteria (step 8). These parameters were all retained for the current runs in this report.

- Surface water samples using only parameters with highly positive coefficients (17 parameters with loadings > 0.5). This run was performed to determine the effect on variance. Although the amount of variance related to PC1 and PC2 increased, the ability to distinguish groups of potential contamination impact were not as distinct. For the current report, the practice of using as large amount of parameters as possible was continued.

- Surface water samples with and without base flow distinguished from high flow samples. These runs were performed to determine differences in impact at high flow and base flow as observed in the scores and evaluate any bias of sampling during high flow. In this current report, all surface water samples are designated as either high flow or base flow samples.

- Surface water samples without edge of field samples. This run was performed to determine the influence of edge of field samples on the results. This run was also performed for the current report.

- Surface water samples without the samples with the highest 22 PC1 scores. This run was performed to determine the influence of samples with high concentrations.

- Surface and groundwater samples with and without additional phosphorus parameters. As above, this run was performed to determine the influence of using three forms of phosphorus.

- Surface and groundwater samples with and without samples with the highest 22 PC1 scores.

- Surface and groundwater samples with the phosphorus (4500PF) results replaced with dissolved phosphorus (6020) and total metals replaced with dissolved metals for geoprobe samples. This replacement provides lower values for the phosphorus and metal concentrations. The geoprobe sample typically had high turbidity (geoprobes are not developed similar to wells) and therefore, total concentrations are elevated. These substitutions were continued for the current report.

- Surface and spring samples only. This was performed to see the scores and influence of springs with observed or potential cattle contamination. This run was also performed for the current report.

As a result of the previous PCA runs, the evaluations for this report also included a series of investigative and sensitivity runs. These various runs are summarized in Table 6.11-7 (see last column for purpose) and discussed in the following paragraphs:

- A series of PCA runs were conducted to evaluate the sensitivity on the water PCA of using total versus dissolved metals concentrations: SW 1 versus SW 2, SW 3 versus SW 4, SW 5 versus SW 6, SW 7 versus SW 8, SW 9 versus SW 10, SW 11 versus SW 12, and SW 13 versus SW 14. These runs were conducted under a variety of other sensitivity conditions (discussed below). In all of these runs, changes in the PCA results were observed to be minor; i.e., the results were similar whether total or dissolved metals were used. Although similar, the PCA runs with total metals did exhibit a generally stronger relationship or ability to characterize waste source signatures in the watershed. This was reasonable because the impacts were expected to be more significant during high flow conditions.

- A series of PCA runs were conducted to evaluate the sensitivity on the water PCA, and on the solids PCA, of allowing missing data in the calculation of PC scores versus not allowing any missing data: SW 3 versus SW 15, SW 16 versus SW 17, SD 1 versus SD 2, SD 3 versus SD 4, and SD 6 versus SD 7. These runs were conducted under a variety of other sensitivity conditions (discussed below). In all of these runs, changes in the PCA were either observed to be minor, or the results were similar between corresponding samples. Although similar, the PCA runs that allowed for relatively larger amounts of missing data did provide relatively more information (more sample PC scores) for purposes of evaluating waste source signatures in the watershed.

- A series of PCA runs were conducted to evaluate the sensitivity on the water PCA of using one phosphorus variable versus using three (possibly redundant) phosphorus variables, in conjunction with the sensitivity of using a single bacteria variable versus using multiple bacteria variables: SW 7 versus SW 8 (one versus three phosphorus variables), SW 9 versus SW 10 and SW 11 versus SW 12 (one versus multiple bacteria variables), and SW 13 versus SW 14 (combination of one versus three phosphorus, and one versus multiple bacteria). In addition, these runs were conducted with total versus dissolved metals (discussed above). The runs using a single bacteria versus multiple bacteria variables was conducted to test the possible impact on the PCA of multiple bacteria all with high concentrations. In all of these runs, changes in the PCA were either observed to be minor, or the results were similar between corresponding samples. Although similar, the PCA runs that included all three forms of phosphorus, and that included multiple bacteria variables, did exhibit a generally stronger relationship or ability to characterize waste source signatures in the watershed.

- A series of PCA runs were conducted to evaluate the sensitivity on the water PCA of including SPLP leachate data and/or edge-of-field data versus not including these data: SW 1 versus SW 3, SW 2 versus SW 4, SW 3 versus SW 5, and SW 4 versus SW 6. These runs were conducted to investigate the relative impact on the PCA of including samples with much higher overall concentrations, i.e., potentially more indicative of poultry and cattle impacts. In all of these runs, including these data generally enhanced the ability to evaluate waste source signatures in the watershed. However, the SPLP samples had a significant impact on the PCA results, essentially overwhelming all other sample results and decreasing the ability to distinguish source impact in ambient surface waters of the IRW. Therefore, these runs indicated that including the SPLP data was not representative of actual source impact conditions in the watershed. Additional PCA runs were conducted to further evaluate differences between SPLP and edge-of-field samples only. These runs: SW 16, SW 17, SW 18, SW 19, SW 20, and SW 21, which were considered more "investigative" in nature, provided further support for excluding the SPLP data in the selection of the most important runs for evaluating source signatures.

- A series of PCA runs were conducted to evaluate the sensitivity or influence on the water PCA of including groundwater and/or spring sample data versus not including these data: SW 3 versus SW 16, SW 17, and SW 18. These runs were conducted to evaluate the relative impact on the PCA of including samples (homeowner groundwater) with much lower overall concentrations. In all of these runs, including these data did not negatively impact the ability to evaluate waste source signatures in the watershed. In certain cases, the inclusion of these data, especially the spring samples, was useful in interpreting or explaining certain apparently anomalous results.

- A series of PCA runs were conducted to evaluate the sensitivity on the solids PCA of including poultry waste and cow manure sample data versus not including these data: SD 1 versus SD 3, and SD 2 versus SD 4. Similar to the SPLP leachate and edge of field water sensitivity runs, these runs were conducted to investigate the relative impact on the solids PCA of including samples with much higher overall concentrations, i.e., potentially more indicative of poultry and cattle impacts. In all of these runs, including these data generally enhanced the ability to evaluate waste source signatures in the watershed.

- Additional solids PCA runs were conducted to evaluate the impact on the PCA of including Lake Tenkiller core samples versus not including these samples: SD 1 versus SD 6, SD 7, and SD 8. In addition, an investigative PCA run using only Lake Tenkiller core samples was conducted: SD 5. All of these runs were used to evaluate whether the core samples could be included in the PCA without loss of information and without biasing the results, due to the fact that the core samples were necessarily analyzed for a smaller set of variables (limited amount of material was available for analysis). The results indicated that including the core samples supplied additional information relevant to the evaluation of waste source signatures in the watershed.

The above sensitivity runs relate to the current PCA runs conducted and discussed in this report. However, in addition to these current runs, numerous sensitivity runs were also conducted during previous, preliminary PCA runs. As discussed above, many of these previous runs were repeated in the current runs and are therefore not discussed specifically in this report. On the other hand, some of these previous runs were not repeated, including, for example, the sensitivity on the water PCA of including arsenic and nickel data versus not including these data.

In summary, the sensitivity analyses indicated that the PCA (as established and conducted in this investigation) proved to be very robust and was insensitive to changes in variables, groupings, or other conditions. The PCA is an appropriate method to identify major sources of contamination in the IRW.

Step 15: State and Document Conclusions

Overall, PCA supports the other lines of evidence previously discussed in this section. Major conclusions from the PCA follow:

- PCA identified two major sources of contamination in the IRW: poultry waste disposal and WWTP discharges. Poultry waste is by far the dominant contamination source in the IRW when compared to other sources. Cattle waste contamination was unique from both poultry waste and WWTP discharges; however, contamination from cattle waste is not dominant in the IRW and only represents a minor source.

The overall conclusions of the PCA evaluation in relation to the hypotheses given in section 6.1 follow:

- Land application of poultry waste affects the chemical and bacterial water and sediment composition of the IRW. The affect is observable in surface water, groundwater and sediments collected from the IRW. This is shown by PCA: a large and distinct group of samples is dominated by poultry waste contamination.

- WWTP discharges into rivers affect the chemical and bacterial water composition of the IRW. The affect is observable in surface waters collected from the IRW. This is shown by PCA: a distinct group of samples is dominated by WWTP discharge.

- Cattle manure deposited in fields and rivers affects the chemical and bacterial composition; however, no dominant impact is observed from cattle waste in the PCA. This is consistent with the mass balances.

12

6.1~~1.3~~ Conclusions

As discussed in Section 6.2, multiple lines of evidence were used to evaluate the sources of contamination in the IRW. The multiple lines of evidence all support that poultry waste disposal by land application is a major source of contamination including phosphorus and bacteria in the IRW. These lines of evidence include the chemical and bacterial composition of major waste sources compared to contamination in the IRW, mass balance calculations showing that poultry waste is a major source of contamination, fate and transport observations for poultry waste contaminants through out the IRW, analyses and detection of a poultry specific biomarker and PCA evaluations showing poultry waste contamination in a dominant source. These lines of evidence can be used to test the hypotheses stated in Section 6.1. The conclusions concerning the hypotheses follow:

- Land application of poultry waste affects the chemical and bacterial water and sediment composition of the IRW and the affect is observable in surface water, groundwater and sediments collected from the IRW. Poultry waste is the dominant source of contamination in the IRW.

- WWTP discharges into rivers affect the chemical and bacterial water composition of the IRW. The affect is observable in surface waters collected from the IRW. The effect is not as large as the effect of poultry waste disposal in the IRW.

- Cattle manure deposited in fields and rivers affects the chemical and bacterial composition; however, no dominant impact is observed from cattle waste in the PCA.

Prepared By:

Roger L. Olsen

May 14, 2008

My hourly rate is $260.00/hour

CDM

Appendix F
Principal Component Scores (PC1 and PC2)
SW3, SW17, SD1 (Varimax), SD6, SW18

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
BS-08:8/23/2005:SW:S:-:-	1.2505508234	3.4776342730
BS-117:9/14/2005:SW:S:-:-	1.4145731216	3.7642780441
BS-208:5/1/2007:SW:S:0:-	1.2706651191	3.3730112723
BS-208:9/1/2005:SW:S:-:-	1.2842496981	3.3293785312
BS-28:8/23/2005:SW:S:-:-	1.2615482519	3.7045307341
BS-35:5/2/2007:SW:S:0:-	1.3127338516	3.8463264796
BS-35:9/22/2005:SW:S:-:-	1.4390547346	3.8050638165
BS-62A:5/1/2007:SW:S:0:-	1.3235155117	3.8627513059
BS-68:5/2/2007:SW:S:0:-	1.2786548761	3.9321464578
BS-HF04:5/1/2007:SW:S:0:-	1.5373208603	5.7424007881
BS-HF22:5/2/2007:SW:S:0:-	1.5205054133	4.8435067403
BS-HF22:8/24/2005:SW:S:-:-	1.5298193012	4.9471816482
BS-REF1:8/30/2005:SW:S:-:-	1.1834742751	3.6878162516
BS-REF2:8/31/2005:SW:S:-:-	1.0000000000	4.3736537929
BS-REF3:9/1/2005:SW:S:-:-	1.1878138051	3.6366685531
EOF01:5/23/2005:SW:S:-:-	4.4818964561	2.1085453867
EOF02:5/23/2005:SW:S:-:-	5.9370147958	1.1444228189
EOF07:5/15/2005:SW:S:-:-	6.6122026030	3.8745314254
EOF07:5/23/2005:SW:S:-:-	1.9457262837	3.6010395810
EOF07:6/5/2005:SW:S:-:-	2.0918148131	4.1529180119
EOF07-222:4/13/2007:SW:S:-:-	2.7874494503	3.4658401378
EOF07-230:4/24/2007:SW:S:-:-	1.9186996988	3.8627136184
EOF07-232:4/24/2007:SW:S:-:-	2.4965378698	2.5951687268
EOF07-259:4/13/2007:SW:S:-:-	2.9523718813	2.6629869615
EOF07-LOR#1:4/24/2007:SW:S:-:-	2.5094688437	3.2889782400
EOF08:5/23/2005:SW:S:-:-	1.7367737895	3.8739167854
EOF09:6/5/2005:SW:S:-:-	2.1355448415	3.8424229413
EOF11:6/5/2005:SW:S:-:-	2.5515570131	3.7794119770
EOF14:6/2/2005:SW:S:-:-	6.1452763456	6.0026497009
EOF15:6/2/2005:SW:S:-:-	2.0695467922	2.6886608693
EOF16:6/5/2005:SW:S:-:-	2.4541289679	2.8462972574
EOF17:6/5/2005:SW:S:-:-	2.4461763465	2.7652669043
EOF18:6/5/2005:SW:S:-:-	1.7060627008	3.5110065633
EOF19:6/5/2005:SW:S:-:-	1.7767845024	3.1690299027
EOF20:6/5/2005:SW:S:-:-	3.9047350807	3.1590242798
EOF21:6/5/2005:SW:S:-:-	2.0071555983	4.4694506339
EOF22:6/5/2005:SW:S:-:-	2.1900239721	5.6294381345
EOF23:6/5/2005:SW:S:-:-	1.8082038278	2.6543803175
EOF24:6/5/2005:SW:S:-:-	2.5251595736	2.5737959056
EOF25:6/5/2005:SW:S:-:-	2.0020758669	2.8586554922
EOF26:6/5/2005:SW:S:-:-	2.0923552659	4.3090244319
EOF27:6/5/2005:SW:S:-:-	3.2536628649	2.6872189287
EOF28:6/5/2005:SW:S:-:-	2.2212633655	3.7220604951
EOF29:6/5/2005:SW:S:-:-	5.7839470681	1.0000000000
EOF30:6/5/2005:SW:S:-:-	2.7131088876	2.4261967744
EOF-CP-1A:3/31/2008:SW:S:-:-	1.6790976667	2.8696764214
EOF-CP-1B:3/31/2008:SW:S:-:-	2.1069060719	3.1842598874
EOF-EOF1:6/17/2006:SW:S:-:-	2.2043450547	2.7396655039
EOF-GF1:3/9/2006:SW:S:-:-	1.9390451662	3.1695746486
EOF-SPREAD002:4/25/2006:SW:S:-:-	2.2095379002	3.8362720455

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	**PC 1**	**PC 2**
EOF-SPREAD007:4/25/2006:SW:S:-:-	2.3944310516	2.6164587927
EOF-SPREAD007:5/4/2006:SW:S:-:-	2.0917590070	3.2856504512
EOF-SPREAD010:5/9/2006:SW:S:-:-	6.2292072162	5.8732848684
EOF-SPREAD017A:5/1/2006:SW:S:-:-	6.9106047198	4.8183414321
EOF-SPREAD023:4/25/2006:SW:S:-:-	8.0983224800	9.4664884801
EOF-SPREAD023:6/18/2006:SW:S:-:-	6.6792993132	9.0800003516
EOF-SPREAD025:5/4/2006:SW:S:-:-	1.6410881542	2.8597515989
EOF-SPREAD025:6/18/2006:SW:S:-:-	1.3823091899	4.0560300153
EOF-SPREAD026:4/25/2006:SW:S:-:-	2.6164814021	2.4778236173
EOF-SPREAD026:4/29/2006:SW:S:-:-	2.0119167769	2.6297763189
EOF-SPREAD029:4/25/2006:SW:S:-:-	3.0251167731	3.0430185766
EOF-SPREAD030:3/31/2006:SW:S:-:-	1.8943017305	4.3404471682
EOF-SPREAD031:4/7/2006:SW:S:-:-	1.6485522237	4.2107841591
EOF-SPREAD036:4/25/2006:SW:S:-:-	3.3045442406	5.2784315208
EOF-SPREAD044:6/18/2006:SW:S:-:-	1.4463286140	4.2562516027
EOF-SPREAD048:5/9/2006:SW:S:-:-	2.4649238601	3.3886498389
EOF-SPREAD052:4/25/2006:SW:S:-:-	1.9109324514	3.4464616261
EOF-SPREAD053B:5/4/2006:SW:S:-:-	1.9423720683	4.0000983128
EOF-SPREAD053E:4/29/2006:SW:S:-:-	7.6490409287	9.6341597904
EOF-SPREAD053G:5/4/2006:SW:S:-:-	2.2926230103	3.3604252366
EOF-SPREAD059:4/29/2006:SW:S:-:-	1.5834908548	3.4890514931
EOF-SPREAD060:4/29/2006:SW:S:-:-	6.8966308258	9.3375530707
EOF-SPREAD064:5/4/2006:SW:S:-:-	5.6665853034	4.1394873574
EOF-SPREAD065:5/4/2006:SW:S:-:-	1.7366661166	3.3158006959
EOF-SPREAD068:6/18/2006:SW:S:-:-	1.6891670420	3.7313342401
EOF-SPREAD071:5/10/2006:SW:S:-:-	1.7557500985	3.2985125182
EOF-SPREAD073E:6/22/2006:SW:S:-:-	1.8682488946	2.9306347323
EOF-ZPEOF001:4/25/2006:SW:S:-:-	3.3824044662	2.4593256724
EOF-ZPEOF030:4/25/2006:SW:S:-:-	1.8605726402	3.4505919464
HFS-02:4/29/2006:SW:S:-:INITIAL	1.3336815355	3.6017497812
HFS-02:5/1/2006:SW:S:-:-	1.3975630432	3.7616557394
HFS-02:5/10/2006:SW:S:-:-	1.3670297221	3.7651688752
HFS-02:5/11/2006:SW:S:-:PEAK	1.3952601855	3.6471265932
HFS-02:5/4/2006:SW:S:-:PEAK	1.4221896304	3.7009750589
HFS-02:5/6/2006:SW:S:-:-	1.3751948984	3.7451841987
HFS-02:6/15/2005:SW:S:-:-	1.3766300126	3.5236622989
HFS-02:6/15/2006:SW:S:-:BF1	1.2310134253	3.6628628516
HFS-02:6/27/2005:SW:S:-:-	1.3685426951	3.5866895243
HFS-02:6/8/2006:SW:S:-:-	1.3471572113	3.6545840876
HFS-02:7/11/2005:SW:S:-:-	1.3140080835	3.5653968545
HFS-02:7/11/2005:SW:S:-:BF1	1.3210744839	3.8163217946
HFS-02:8/1/2006:SW:S:-:BF2	1.2832458648	3.5677550442
HFS-02:8/27/2005:SW:S:-:BF2	1.3053131863	3.6177405215
HFS-02:9/16/2005:SW:S:-:-	1.3025218188	3.5529792741
HFS-04:3/10/2006:SW:S:-:-	1.7294026164	6.1749557317
HFS-04:3/10/2006:SW:S:-:PEAK	1.7047056185	6.0003051089
HFS-04:4/26/2006:SW:S:-:-	1.7479152448	6.0097489461
HFS-04:4/30/2006:SW:S:-:-	1.5769140279	4.7911150554
HFS-04:4/4/2006:SW:S:-:-	1.7132714950	6.1613180043
HFS-04:4/8/2006:SW:S:-:-	1.6237631640	5.8401397293

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
HFS-04:5/11/2006:SW:S:-:EVENTB	1.6523831179	4.2465910379
HFS-04:5/25/2005:SW:S:-:1	1.5796054835	4.8074329709
HFS-04:5/25/2005:SW:S:-:14	1.5428687955	4.6087816501
HFS-04:5/25/2005:SW:S:-:7	1.5960202980	4.6974142414
HFS-04:5/31/2006:SW:S:-:-	1.5422598207	4.3106420338
HFS-04:5/31/2006:SW:S:-:PEAK	1.5631703583	4.3022649198
HFS-04:5/6/2006:SW:S:-:-	1.5611978850	5.1430475680
HFS-04:5/9/2006:SW:S:-:EVENTA	1.5536244439	4.9690523656
HFS-04:6/15/2006:SW:S:-:BF1	1.6679145833	7.0249588568
HFS-04:6/22/2005:SW:S:-:A	1.7299210364	6.0451069558
HFS-04:6/22/2005:SW:S:-:B	1.5612221847	5.2987178592
HFS-04:7/13/2005:SW:S:-:BF1	1.6906856886	6.2546574969
HFS-04:7/27/2005:SW:S:-:-	1.4824832316	4.6803488984
HFS-04:8/1/2006:SW:S:-:BF2	1.7648148212	7.6169367661
HFS-04:8/20/2005:SW:S:-:A	1.5519810795	5.6864181278
HFS-04:8/20/2005:SW:S:-:B	1.6106163038	5.8377554487
HFS-04:8/27/2005:SW:S:-:BF2	1.7034251577	6.8308590302
HFS-04:9/28/2005:SW:S:-:-	1.6934113148	7.2890708901
HFS-05:3/22/2006:SW:S:-:-	2.2177295516	4.4695680795
HFS-05:3/22/2006:SW:S:-:PEAK	1.4320257293	4.8427370588
HFS-05:4/26/2006:SW:S:-:-	1.4576405494	4.1828231428
HFS-05:4/26/2006:SW:S:-:PEAK	1.4505168803	3.8818363971
HFS-05:4/29/2006:SW:S:-:PEAK	1.4025299517	4.1533207761
HFS-05:4/3/2006:SW:S:-:-	1.4180236142	4.3680647104
HFS-05:4/30/2006:SW:S:-:-	1.5129514753	4.1930336496
HFS-05:5/10/2006:SW:S:-:EVENTA	1.9543268266	6.2327961046
HFS-05:5/11/2006:SW:S:-:EVENTB	1.6102554410	4.0026939731
HFS-05:5/31/2006:SW:S:-:-	1.3941130663	4.1155146166
HFS-05:5/6/2006:SW:S:-:-	1.5263280989	4.2162864209
HFS-05:6/15/2006:SW:S:-:BF1	1.2698065225	4.3901885967
HFS-05:6/27/2005:SW:S:-:A	1.4127233052	3.8270667820
HFS-05:6/27/2005:SW:S:-:B	1.5370088483	3.7042563561
HFS-05:6/7/2005:SW:S:-:A	1.5181727976	4.1273117643
HFS-05:6/7/2005:SW:S:-:B	1.7516272030	3.6160307302
HFS-05:6/7/2006:SW:S:-:-	1.3351031336	4.1528961708
HFS-05:7/12/2005:SW:S:-:BF1	1.3517998946	4.0285537502
HFS-05:7/7/2005:SW:S:-:-	1.3607729080	4.0152166349
HFS-05:8/1/2006:SW:S:56:BF2	1.3911795413	3.8926447187
HFS-05:8/29/2005:SW:S:-:BF2	1.2436923686	4.3444649430
HFS-05:9/13/2005:SW:S:-:A	1.4077623279	3.5030224684
HFS-05:9/15/2005:SW:S:-:B	1.3776490729	3.9205220883
HFS-05:9/28/2005:SW:S:-:-	1.2997874315	3.6118752797
HFS-08:6/15/2005:SW:S:-:A	1.4879581475	3.6455951838
HFS-08:6/15/2005:SW:S:-:B	1.4437547434	3.3356119034
HFS-08:7/12/2005:SW:S:-:-	1.4170108637	3.2490046007
HFS-08:7/13/2005:SW:S:-:BF1	1.3515417748	3.3500626455
HFS-08:8/28/2005:SW:S:-:BF2	1.2574581397	3.5743880110
HFS-14:3/10/2006:SW:S:-:-	1.3964472981	3.7199055350
HFS-14:3/10/2006:SW:S:-:PEAK	1.6036982808	4.0134250477
HFS-14:4/25/2006:SW:S:-:-	1.4847393708	3.8641406911

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
HFS-14:4/25/2006:SW:S:-:PEAK	1.6079019998	3.8553521697
HFS-14:4/7/2006:SW:S:-:-	1.3430419633	3.7281763890
HFS-14:4/7/2006:SW:S:-:PEAK	1.3601402608	3.8536445155
HFS-14:5/10/2006:SW:S:-:-	1.4234763291	3.8212089011
HFS-14:5/11/2006:SW:S:-:PEAK	1.6446447667	4.1570079551
HFS-14:5/31/2006:SW:S:-:-	1.8405647786	4.3625485316
HFS-14:6/15/2005:SW:S:-:A	1.3694239564	3.4206232614
HFS-14:6/15/2005:SW:S:-:B	1.3020895831	3.3818423647
HFS-14:6/15/2006:SW:S:-:BF1	1.1818754636	3.5425163162
HFS-14:6/5/2006:SW:S:-:-	1.3858003647	3.9222152169
HFS-14:6/9/2005:SW:S:-:-	1.2901941108	3.4257222824
HFS-14:7/12/2005:SW:S:-:BF1	1.3616815638	3.2736216087
HFS-14:7/23/2005:SW:S:-:-	1.2310938454	3.2966078939
HFS-14:8/1/2006:SW:S:-:BF2	1.2572549054	3.4821004004
HFS-14:8/27/2005:SW:S:-:BF2	1.2288536135	3.5283839235
HFS-16:3/10/2006:SW:S:-:-	1.5273309318	3.6445465527
HFS-16:4/25/2006:SW:S:-:-	1.4936446225	3.4308003246
HFS-16:5/1/2006:SW:S:-:PEAK	1.3134058758	3.8249566721
HFS-16:5/2/2006:SW:S:-:-	1.3008076368	3.8722290344
HFS-16:5/31/2006:SW:S:-:-	1.4046815062	3.4487318100
HFS-16:5/6/2006:SW:S:-:A	1.3181566054	3.9136746480
HFS-16:5/7/2006:SW:S:-:B	1.4002381586	4.4813123357
HFS-16:6/15/2005:SW:S:-:-	1.5007374150	3.7453106111
HFS-16:6/15/2006:SW:S:-:BF1	1.3455762674	3.7688587297
HFS-16:6/27/2005:SW:S:-:A	1.4310342060	3.7049862589
HFS-16:7/11/2005:SW:S:-:BF1	1.8174216798	3.7309099056
HFS-16:7/13/2005:SW:S:-:-	1.3691011583	3.8564536047
HFS-16:8/14/2005:SW:S:-:-	1.5186468550	3.7711223294
HFS-16:8/27/2005:SW:S:-:BF2	1.3564378635	4.2749764045
HFS-16:9/15/2005:SW:S:-:-	1.4622660450	3.8259137915
HFS-20:3/10/2006:SW:S:-:-	1.3294432254	3.8662752400
HFS-20:4/26/2006:SW:S:-:-	1.3090851934	3.7924872520
HFS-20:4/3/2006:SW:S:-:-	1.3029038024	3.9198115606
HFS-20:4/3/2006:SW:S:-:PEAK	1.3039006006	3.8930443048
HFS-20:4/30/2006:SW:S:-:-	1.3109480235	4.1136454171
HFS-20:4/7/2006:SW:S:-:-	1.1992472148	4.1136901571
HFS-20:5/10/2006:SW:S:-:EVENTB	1.6111287409	3.3904302434
HFS-20:5/31/2006:SW:S:-:EVENTA	1.3586301041	3.8132461695
HFS-20:5/31/2006:SW:S:-:EVENTB	1.3686562474	3.7244443515
HFS-20:5/4/2006:SW:S:-:LEADINGEDGE	1.3471085901	4.0936229696
HFS-20:5/6/2006:SW:S:-:PLATEAU	1.3151443018	4.0407525927
HFS-20:5/9/2006:SW:S:-:EVENTA	1.6014585952	3.3701251847
HFS-20:6/15/2005:SW:S:-:A	1.3840496833	3.6544778068
HFS-20:6/15/2005:SW:S:-:B	1.5475382535	3.6350921418
HFS-20:6/15/2006:SW:S:-:BF1	1.1941895907	3.9103049919
HFS-20:7/10/2005:SW:S:-:-	1.2094224551	4.1100130772
HFS-20:7/13/2005:SW:S:-:BF1	1.3290078291	4.1071933940
HFS-20:7/26/2005:SW:S:-:-	1.5578477616	3.8356042182
HFS-20:7/7/2005:SW:S:-:A	1.3198684135	4.0609285314
HFS-20:7/7/2005:SW:S:-:B	1.3467999829	4.3134574332

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
HFS-20:8/1/2006:SW:S:-:BF2	1.4635936370	3.8607187492
HFS-20:8/14/2005:SW:S:-:-	1.1920633657	4.2677947760
HFS-20:8/27/2005:SW:S:-:BF2	1.2626557429	4.4167147044
HFS-20:9/16/2005:SW:S:-:-	1.4881550360	3.5101013146
HFS-21:3/21/2006:SW:S:-:-	1.4046753079	3.4112190516
HFS-21:3/9/2006:SW:S:-:-	1.4623037062	3.4208834975
HFS-21:4/26/2006:SW:S:-:-	1.4950326783	3.2915128407
HFS-21:4/26/2006:SW:S:-:PEAK	1.4465379294	3.1304432638
HFS-21:4/29/2006:SW:S:-:PEAK	1.4995208751	3.3007951203
HFS-21:4/3/2006:SW:S:-:-	1.4483626012	3.3247606009
HFS-21:4/30/2006:SW:S:-:-	1.4742108235	3.2846064632
HFS-21:4/6/2006:SW:S:-:A	1.4527998735	4.0354892653
HFS-21:4/7/2006:SW:S:-:B	1.5308768727	3.4487527492
HFS-21:5/10/2006:SW:S:-:EVENTB	1.3603582539	3.8294218129
HFS-21:5/31/2006:SW:S:-:-	1.4667864514	3.3153728776
HFS-21:5/4/2006:SW:S:-:PEAK	1.5661032288	3.6942794956
HFS-21:5/5/2006:SW:S:-:-	1.5259651239	3.1690231937
HFS-21:5/9/2006:SW:S:-:EVENTA	1.3159738362	3.8940525775
HFS-21:6/15/2006:SW:S:-:BF1	1.4806258043	3.5414501595
HFS-21:6/7/2005:SW:S:-:-	1.5172926371	3.2912401175
HFS-21:7/10/2005:SW:S:-:-	1.4939670987	3.5614819709
HFS-21:7/13/2005:SW:S:-:BF1	1.8318387424	4.1436310877
HFS-21:7/7/2005:SW:S:-:-	1.5002780019	3.4143547915
HFS-21:8/14/2005:SW:S:-:-	1.6128629441	3.8478197276
HFS-21:8/20/2005:SW:S:-:-	1.3642109807	3.7088615145
HFS-21:9/16/2005:SW:S:-:-	1.5866524532	3.9401775282
HFS-21:9/28/2005:SW:S:-:-	1.5713720514	3.6764004987
HFS-22:3/22/2006:SW:S:-:-	1.6210168217	4.8385850501
HFS-22:3/22/2006:SW:S:-:PEAK	1.6813117049	5.0619170404
HFS-22:4/27/2006:SW:S:-:-	1.8117528449	5.2359512290
HFS-22:4/27/2006:SW:S:-:PEAK	1.8485254608	5.4354187579
HFS-22:4/30/2006:SW:S:-:-	1.5841753456	3.6894033799
HFS-22:4/4/2006:SW:S:-:-	1.5910787556	4.8197236338
HFS-22:4/4/2006:SW:S:-:PEAK	1.6226334148	4.9753667809
HFS-22:5/10/2006:SW:S:-:EVENTA	1.9975495498	5.1436172032
HFS-22:5/11/2006:SW:S:-:EVENTB	1.5814300541	3.5699574840
HFS-22:5/31/2006:SW:S:-:EVENTB	1.4920579723	3.8821848908
HFS-22:5/6/2006:SW:S:-:-	1.5694384988	3.7079634469
HFS-22:6/15/2005:SW:S:-:A	1.6445375795	5.0170734261
HFS-22:6/15/2005:SW:S:-:B	1.6559916797	4.8836970220
HFS-22:6/15/2006:SW:S:-:BF1	1.5089488956	5.0237771165
HFS-22:6/7/2005:SW:S:-:-	1.6379439857	5.1052030260
HFS-22:7/12/2005:SW:S:-:BF1	1.5680496634	4.9225836088
HFS-22:8/1/2006:SW:S:1:BF2	1.7187944091	5.6284339413
HFS-22:8/28/2005:SW:S:-:BF2	1.5695327553	5.5861181548
HFS-22:9/15/2005:SW:S:-:-	1.2424179717	3.9617321673
HFS-23:3/22/2006:SW:S:-:-	1.4752815840	4.0464655675
HFS-23:4/27/2006:SW:S:-:-	1.5695809209	4.1223543811
HFS-23:4/27/2006:SW:S:-:PEAK	1.5656955141	3.8873472617
HFS-23:4/30/2006:SW:S:-:-	1.4751440077	3.8938825433

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
HFS-23:4/4/2006:SW:S:-:-	1.4615499889	4.0283008808
HFS-23:4/8/2006:SW:S:-:-	1.4701025782	4.1470567106
HFS-23:5/10/2006:SW:S:-:PEAK	1.5221467408	3.4954990666
HFS-23:5/11/2006:SW:S:-:-	1.4671700313	3.4727709609
HFS-23:5/31/2006:SW:S:-:-	1.5563314280	3.7760427408
HFS-23:5/4/2006:SW:S:-:G	1.5685544709	3.2966277066
HFS-23:6/15/2005:SW:S:-:-	1.5500439268	3.8936250010
HFS-23:6/15/2006:SW:S:-:BF1	1.5875339257	4.4358461812
HFS-23:6/27/2005:SW:S:-:-	1.5175876283	4.2141213304
HFS-23:6/7/2005:SW:S:-:-	1.5532412254	4.1229397903
HFS-23:6/7/2006:SW:S:-:-	1.4276593208	4.0417177560
HFS-23:7/12/2005:SW:S:-:BF1	1.7024663882	3.8825989762
HFS-23:7/16/2005:SW:S:-:-	1.5011937751	3.9247174626
HFS-23:7/23/2005:SW:S:-:-	1.5567613556	3.8023059341
HFS-23:8/1/2006:SW:S:56:BF2	1.5043715072	4.1697450097
HFS-23:8/14/2005:SW:S:-:A	1.5605293026	4.5382627199
HFS-23:8/14/2005:SW:S:-:B	1.5773309164	4.5547670985
HFS-23:8/28/2005:SW:S:-:BF2	1.4932183088	4.6639783704
HFS-23:9/16/2005:SW:S:-:-	1.6826362847	3.9504751706
HFS-23:9/25/2005:SW:S:-:-	1.8051470082	3.9824481473
HFS-26:7/13/2005:SW:S:-:BF1	1.3363683756	3.5008918032
HFS-26:7/7/2005:SW:S:-:-	1.3245373887	3.4907980890
HFS-26:8/27/2005:SW:S:-:BF2	1.2263305494	3.7997629997
HFS-28A:4/25/2006:SW:S:-:-	1.2261277508	3.3379951438
HFS-28A:4/26/2006:SW:S:-:TAIL	1.2832180888	3.3136781321
HFS-28A:4/29/2006:SW:S:-:RISINGLIMB	1.2688142385	3.3404852720
HFS-28A:4/30/2006:SW:S:-:-	1.2697066829	3.3956807730
HFS-28A:5/10/2006:SW:S:-:-	1.3478006903	3.2925342541
HFS-28A:5/4/2006:SW:S:-:LEADINGEDGE	1.2958982881	3.3617816510
HFS-28A:5/5/2006:SW:S:-:-	1.2702497475	3.3405244935
HFS-28A:6/15/2006:SW:S:-:BF1	1.1985009841	3.3339267619
HFS-28A:6/7/2006:SW:S:-:-	1.2564057640	3.0247613619
HFS-28A:8/1/2006:SW:S:-:BF2	1.2488777394	3.3940878448
HFS-28A:8/28/2005:SW:S:-:BF2	1.1998353268	3.6395292148
HFS-29:3/10/2006:SW:S:-:-	1.3544725515	3.4407299958
HFS-29:4/24/2006:SW:S:-:-	1.3872731454	3.5173809006
HFS-29:4/29/2006:SW:S:-:-	1.3363992218	3.6549366899
HFS-29:4/29/2006:SW:S:-:PEAK	1.4268034938	3.4607376720
HFS-29:4/7/2006:SW:S:-:-	1.3511503156	3.5269150141
HFS-29:4/7/2006:SW:S:-:PEAK	1.3047685366	3.4437483081
HFS-29:5/10/2006:SW:S:-:EVENTA	1.4190801384	3.7991069497
HFS-29:5/10/2006:SW:S:-:EVENTB	1.4348814087	3.6892910743
HFS-29:5/5/2006:SW:S:-:-	1.4565995989	4.1680300941
HFS-29:6/15/2006:SW:S:-:BF1	1.3143633745	3.6932205659
HFS-29:8/1/2006:SW:S:-:BF2	1.3411996069	3.5030194242
HFS-30:4/24/2006:SW:S:-:-	1.3000413417	3.3701940305
HFS-30:4/26/2006:SW:S:-:TAIL	1.2769687838	3.4360580345
HFS-30:4/29/2006:SW:S:-:PEAK	1.2846080045	3.4304887596
HFS-30:4/30/2006:SW:S:-:-	1.2889505956	3.4056169407
HFS-30:5/11/2006:SW:S:-:-	1.2924425628	3.0375469601

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
HFS-30:5/3/2007:SW:S:0:-	1.2273361471	3.1717963779
HFS-30:5/31/2006:SW:S:-:-	1.2951043171	3.2140157362
HFS-30:5/4/2006:SW:S:-:-	1.3769679579	3.0355864282
HFS-30:6/15/2006:SW:S:-:BF1	1.1919231735	3.4129818830
HFS-30:8/1/2006:SW:S:-:BF2	1.2220531165	3.4421369288
LincolnWWTP:4/2/2008:SW:S:-:-	1.4340495414	3.7321837074
LK-01:7/12/2005:SW:S:1:-	1.4286330751	3.2193905891
LK-01:7/12/2005:SW:S:14:-	1.4282510148	3.4432763967
LK-01:7/12/2005:SW:S:18:-	1.4181527554	3.4587571154
LK-01:7/12/2005:SW:S:25:-	1.4576849464	3.3893793981
LK-01:8/24/2005:SW:S:1:-	1.3604275457	3.3183082623
LK-01:8/24/2005:SW:S:14:-	1.4132102395	3.4490349545
LK-01:9/26/2006:SW:S:0:-	1.3148184288	3.5918798800
LK-01:9/26/2006:SW:S:16:-	1.3544417415	3.5609371447
LK-01:9/26/2006:SW:S:20:-	1.4390385590	3.5256838468
LK-02:7/12/2005:SW:S:1:-	1.4103531989	3.2755475357
LK-02:7/12/2005:SW:S:10:-	1.4742406477	3.3995324868
LK-02:7/12/2005:SW:S:22:-	1.5839841761	3.3740194639
LK-02:8/23/2005:SW:S:1:-	1.4285556540	3.2972145825
LK-02:8/23/2005:SW:S:10:-	1.4315290534	3.3508471484
LK-02:8/23/2005:SW:S:21:-	1.6982184218	3.3901343263
LK-02:9/26/2006:SW:S:0:-	1.3121602373	3.6372371042
LK-02:9/26/2006:SW:S:16:-	1.3804800982	3.5434614843
LK-02:9/26/2006:SW:S:19:-	1.7295670578	3.3654889630
LK-03:7/12/2005:SW:S:1:-	1.4748569619	3.3412232294
LK-03:7/12/2005:SW:S:3:-	1.4715279525	3.3049116907
LK-03:7/12/2005:SW:S:6:-	1.4563299872	3.3724475328
LK-03:8/23/2005:SW:S:1:-	1.3923983105	3.4217026581
LK-03:8/23/2005:SW:S:3:-	1.4160119231	3.4428948292
LK-03:8/23/2005:SW:S:4:-	1.4304349143	3.4190206454
LK-03:9/26/2006:SW:S:0:-	1.3262227180	3.8058229078
LK-04:8/23/2005:SW:S:1:-	1.4706112001	3.8392974362
LK-04:8/23/2005:SW:S:2:-	1.4932841511	3.8321521744
LK-04:8/23/2005:SW:S:3:-	1.6467507401	3.7633709573
LK-04:9/26/2006:SW:S:0:-	1.4322300551	3.7202993007
RBS-0000019:8/14/2006:SW:S:0:-	2.0748179873	3.0967299457
RBS-0000028:8/14/2006:SW:S:0:-	1.3251399583	3.6802678344
RBS-0000031:8/16/2006:SW:S:0:-	1.4470772204	5.8888790653
RBS-0000043:8/11/2006:SW:S:0:-	1.3631347084	4.4421413318
RBS-0000057:8/11/2006:SW:S:0:-	1.3844170650	4.1370643338
RBS-0000075:8/10/2006:SW:S:0:-	1.5334567012	6.9994862074
RBS-0000086:8/10/2006:SW:S:0:-	1.3664415689	4.0228728106
RBS-0000109:8/9/2006:SW:S:0:-	1.5748554014	5.0532138858
RBS-0000120:8/9/2006:SW:S:0:-	1.3780972592	5.5074857411
RBS-0000121:8/10/2006:SW:S:0:-	1.6414809101	7.8983032640
RBS-0000137:8/11/2006:SW:S:0:-	1.3592308148	4.1429157138
RBS-0000148:8/10/2006:SW:S:0:-	1.2039568830	3.8145936429
RBS-0000150:8/10/2006:SW:S:0:-	1.2318979737	3.7599789381
RBS-0000225:8/11/2006:SW:S:0:-	1.6475201353	4.8522541444
RBS-0000246:8/9/2006:SW:S:0:-	1.5850663362	4.8151650043

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
RBS-0000286:8/9/2006:SW:S:0:-	1.8913900219	3.7715141593
RBS-0000312:8/9/2006:SW:S:0:-	1.2374848775	3.6051911179
RBS-0000336:8/16/2006:SW:S:0:-	1.4101674851	3.6229161649
RBS-0000340:8/15/2006:SW:S:0:-	1.3160397667	4.0525708524
RBS-0000344:8/16/2006:SW:S:0:-	1.2135578460	3.9027877990
RBS-0000345:8/11/2006:SW:S:0:-	1.6777824170	8.0879886638
RBS-0000349:8/11/2006:SW:S:0:-	1.5966128844	6.4681754549
RBS-0000350:8/16/2006:SW:S:0:-	1.4339525565	5.8764149768
RBS-0000395:8/10/2006:SW:S:0:-	1.2431950768	3.8717243074
RBS-0000548:8/15/2006:SW:S:0:-	1.2513401225	3.5750128997
RBS-0000574:8/10/2006:SW:S:0:-	1.4882282937	4.0293605941
RBS-0000577:8/10/2006:SW:S:0:-	1.4424297706	4.1783520334
RBS-0000578:8/10/2006:SW:S:0:-	1.3581186422	4.0013116585
RBS-0000625:8/10/2006:SW:S:0:-	1.3229373421	3.8085145848
RBS-0000630:8/11/2006:SW:S:0:-	1.6215165533	3.5363092996
RBS-0000662:8/15/2006:SW:S:0:-	1.2412426727	3.6767831785
RBS-0000704:8/8/2006:SW:S:0:-	1.2327183154	3.8271762887
RBS-0000706:8/9/2006:SW:S:0:-	1.2428652175	3.5298625978
RBS-0000770:8/11/2006:SW:S:0:-	1.2866354869	3.3826102196
RBS-0000901:8/9/2006:SW:S:0:-	1.4742313283	5.7705945208
RBS-0010003:8/14/2006:SW:S:0:-	1.2327097928	3.4078023599
RBS-0010004:8/14/2006:SW:S:0:-	1.3026607712	3.2331927145
RBS-7198000:8/9/2006:SW:S:0:-	1.3764925577	3.7500831365
RogersWWTP:4/1/2008:SW:S:-:-	1.4719381082	5.0725117220
RS-1:6/13/2006:SW:S:0:-	1.2819939384	3.5709975400
RS-1:7/11/2005:SW:S:0:-	1.3592721522	3.4958632654
RS-1:7/12/2006:SW:S:0:-	1.2237944761	3.4616031276
RS-1:8/24/2005:SW:S:0:-	1.2991798191	3.5157480451
RS-1:8/8/2006:SW:S:0:-	1.2755893718	3.5352109619
RS-1:9/25/2006:SW:S:0:-	1.2716714425	3.6710990116
RS-10004:5/21/2007:SW:S:0:-	1.2449958730	3.1202318267
RS-109:5/2/2007:SW:S:0:-	1.3348387157	4.1474929030
RS-122:5/3/2007:SW:S:0:-	1.3529148275	4.8198087091
RS-133:5/21/2007:SW:S:0:-	1.3545894908	3.6035412835
RS-150:5/3/2007:SW:S:0:-	1.2165880961	3.6325584699
RS-160:5/3/2007:SW:S:0:-	1.2597624218	3.6087469539
RS-2:6/13/2006:SW:S:0:-	1.2480185853	3.4622912374
RS-2:7/11/2005:SW:S:0:-	1.3736662323	3.3560475900
RS-2:7/12/2006:SW:S:0:-	1.2871077374	3.3643801023
RS-2:8/10/2006:SW:S:0:-	1.2928123218	3.4130979912
RS-2:8/24/2005:SW:S:0:-	1.3343907087	3.3997627875
RS-2:9/25/2006:SW:S:0:-	1.3224423295	3.6694023752
RS-233:5/21/2007:SW:S:0:-	1.4025211719	4.0285332786
RS-297:5/21/2007:SW:S:0:-	1.3016844334	4.6893002992
RS-3:6/13/2006:SW:S:0:-	1.3456268636	4.1529832842
RS-3:7/11/2005:SW:S:0:-	1.4255464433	3.8927636351
RS-3:7/12/2006:SW:S:0:-	1.3177557534	4.0126856051
RS-3:8/10/2006:SW:S:0:-	1.3861774433	4.2803674363
RS-3:8/24/2005:SW:S:0:-	1.4093107043	4.0920059745
RS-3:9/25/2006:SW:S:0:-	1.5752410411	3.8909233143

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
RS-312:5/2/2007:SW:S:0:-	1.2260656453	3.4344499313
RS-336:5/10/2007:SW:S:0:-	1.3969586373	3.7749848475
RS-386:5/2/2007:SW:S:0:-	1.3380075734	4.0217141064
RS-399:5/2/2007:SW:S:0:-	1.4989781783	3.7564506694
RS-402:5/21/2007:SW:S:0:-	1.2988213989	3.0500477213
RS-43:5/21/2007:SW:S:0:-	1.3262869865	4.1044482448
RS-433A:5/2/2007:SW:S:0:-	1.2786501205	3.9639694759
RS-578:5/2/2007:SW:S:0:-	1.3022140797	3.8250842056
RS-667:5/1/2007:SW:S:0:-	1.1853333043	3.5150162348
RS-682:5/2/2007:SW:S:0:-	1.2704106832	3.7612554733
RS-696:5/21/2007:SW:S:0:-	1.2761904957	3.4265929021
RS-704:5/2/2007:SW:S:0:-	1.2049289530	3.5119274434
RS-72:5/21/2007:SW:S:0:-	1.3527576177	5.0012755408
RS-728:5/2/2007:SW:S:0:-	1.3597720855	4.3908185761
RS-75:5/8/2007:SW:S:0:-	1.3771127205	5.2543623550
RS-757:5/1/2007:SW:S:0:-	1.3035358203	4.0826083397
RS-770:5/1/2007:SW:S:0:-	1.2570000591	3.3042656299
RS-793:5/3/2007:SW:S:0:-	1.3824249692	4.4569071974
RS-795:5/1/2007:SW:S:0:-	1.3318605296	3.6187100729
RS-902:5/1/2007:SW:S:0:-	1.3349616612	4.0331777903
RS-97:5/21/2007:SW:S:0:-	1.2932831106	4.2077943608
RS-BALLARD:5/5/2006:SW:S:0:-	1.4830087377	3.3464430509
RS-FLYCREEK:5/5/2006:SW:S:0:-	1.4147102484	3.4948643608
RS-ILLINOISRIVER:4/30/2006:SW:S:0:-	1.6843318744	3.5956237357
RS-LOC:4/7/2006:SW:S:0:-	1.6099704644	3.8782400142
RS-OSAGE:5/5/2006:SW:S:0:-	1.3557362954	4.1927509547
RS-PRICECREEK:4/29/2006:SW:S:0:-	1.5272085960	3.2896653457
RS-TYNER:5/5/2006:SW:S:0:-	1.2713160066	3.4389269998
SiloamWWTP:3/31/2008:SW:S:-:-	1.9048742286	6.3564874012
SN-SBC2:4/25/2007:SW:S:0:-	3.4206332775	5.0931703900
SpringdaleWWTP:3/31/2008:SW:S:-:-	1.5783295213	5.9463631151
SSA01:5/14/2005:SW:S:-:-	2.1592172130	4.8248505168
USGS-07195500:10/15/2007:SW:S:-:-	1.3302879661	4.2936201115
USGS-07195500:10/2/2006:SW:S:-:-	1.3693893265	4.2651938366
USGS-07195500:10/4/2007:SW:S:-:-	1.4432706566	3.6788487618
USGS-07195500:12/12/2007:SW:S:-:-	1.3277515426	4.0972479981
USGS-07195500:12/6/2007:SW:S:-:-	1.2832603455	4.5468006208
USGS-07195500:12/7/2006:SW:S:-:-	1.3242302125	4.0467820222
USGS-07195500:2/7/2007:SW:S:-:-	1.2996766185	4.1500373349
USGS-07195500:3/10/2006:SW:S:-:-	1.3672324127	4.8186938312
USGS-07195500:3/9/2006:SW:S:-:-	1.3740795658	4.6642885169
USGS-07195500:4/12/2006:SW:S:-:-	1.3392788561	4.4148580788
USGS-07195500:4/26/2006:SW:S:-:-	1.3897610270	4.7843288510
USGS-07195500:4/26/2007:SW:S:-:-	1.3545519644	3.9563557215
USGS-07195500:4/3/2006:SW:S:-:-	1.3426260772	4.6012609710
USGS-07195500:4/30/2006:SW:S:-:-	1.5480019344	3.7115964129
USGS-07195500:4/9/2007:SW:S:-:-	1.3001506250	4.3431707287
USGS-07195500:5/23/2005:SW:S:-:-	1.3463166786	4.0967911108
USGS-07195500:5/5/2006:SW:S:-:-	1.9792352864	3.2734188589
USGS-07195500:6/12/2007:SW:S:-:-	1.4064401689	3.8290709179

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
USGS-07195500:6/28/2007:SW:S:-:-	1.3365080293	4.1779312245
USGS-07195500:6/4/2007:SW:S:-:-	1.3624425418	3.8974993810
USGS-07195500:7/12/2005:SW:S:-:-	1.3241935188	4.3199113109
USGS-07195500:8/23/2005:SW:S:-:-	1.3715948797	4.2718336915
USGS-07195500:8/6/2007:SW:S:-:-	1.4136419620	4.3064318360
USGS-07195500:9/15/2005:SW:S:-:-	1.3397182919	4.5310108251
USGS-07196000:1/24/2007:SW:S:-:-	1.2996231643	3.7707601174
USGS-07196000:1/9/2008:SW:S:-:-	1.3661013136	3.6253801394
USGS-07196000:10/1/2007:SW:S:-:-	1.3323335155	4.3721585165
USGS-07196000:10/3/2007:SW:S:-:-	1.4207725168	3.8452283169
USGS-07196000:10/5/2006:SW:S:-:-	1.3509287625	4.4781228876
USGS-07196000:12/12/2007:SW:S:-:-	1.3348394544	3.9708269147
USGS-07196000:12/3/2007:SW:S:-:-	1.3681501094	4.4758850291
USGS-07196000:12/6/2006:SW:S:-:-	1.3391070147	4.1162080062
USGS-07196000:2/4/2008:SW:S:-:-	1.3497336502	4.1953218733
USGS-07196000:2/5/2007:SW:S:-:-	1.3064847341	4.0059905777
USGS-07196000:4/12/2006:SW:S:-:-	1.3285749005	4.6326447190
USGS-07196000:4/2/2007:SW:S:-:-	1.3278887008	4.2250480590
USGS-07196000:4/29/2006:SW:S:-:-	1.4002493412	4.5524119421
USGS-07196000:4/30/2006:SW:S:-:-	1.4252644858	4.5277410917
USGS-07196000:5/11/2006:SW:S:-:-	1.3294911948	4.1850565462
USGS-07196000:5/5/2006:SW:S:-:-	1.3903275178	4.3307323773
USGS-07196000:5/8/2007:SW:S:-:-	1.3224466394	3.8794720777
USGS-07196000:6/12/2007:SW:S:-:-	1.3617418749	3.8145482010
USGS-07196000:6/14/2006:SW:S:-:-	1.3335105364	4.4104822299
USGS-07196000:6/5/2007:SW:S:-:-	1.3272393463	4.1047546114
USGS-07196000:7/12/2005:SW:S:-:-	1.3436963139	4.1263227370
USGS-07196000:7/13/2006:SW:S:-:-	1.3629763144	4.4841907517
USGS-07196000:8/24/2005:SW:S:-:-	1.3332043605	4.2503662407
USGS-07196000:8/7/2007:SW:S:-:-	1.4199992135	4.2298714686
USGS-07196090:10/15/2007:SW:S:-:-	1.3030774981	4.1392773986
USGS-07196090:10/4/2007:SW:S:-:-	1.4858779452	3.5193962516
USGS-07196090:10/5/2006:SW:S:-:-	1.3272019620	4.2190397463
USGS-07196090:12/12/2007:SW:S:-:-	1.3064727208	4.1730161691
USGS-07196090:12/6/2007:SW:S:-:-	1.2724830682	4.4182083458
USGS-07196090:12/7/2006:SW:S:-:-	1.3125095174	3.9908729269
USGS-07196090:2/5/2008:SW:S:-:-	1.2980799771	4.2782585744
USGS-07196090:2/7/2007:SW:S:-:-	1.2866830195	4.0302967557
USGS-07196090:3/10/2006:SW:S:-:-	1.2957088057	4.5779387496
USGS-07196090:4/20/2006:SW:S:-:-	1.3191018579	4.4718411360
USGS-07196090:4/26/2006:SW:S:-:-	1.3480316806	4.4653388131
USGS-07196090:4/26/2007:SW:S:-:-	1.3017399357	4.1675352233
USGS-07196090:4/3/2006:SW:S:-:-	1.3078955183	4.5008445614
USGS-07196090:4/30/2006:SW:S:-:-	1.5349683379	4.0109999141
USGS-07196090:4/9/2007:SW:S:-:-	1.2861990238	4.1371131810
USGS-07196090:5/23/2005:SW:S:-:-	1.3434405747	4.0305589986
USGS-07196090:5/5/2006:SW:S:-:-	1.7704431234	3.7995480544
USGS-07196090:6/12/2007:SW:S:-:-	1.3891988369	3.6369411326
USGS-07196090:6/14/2006:SW:S:-:-	1.3151134661	4.2711980268
USGS-07196090:6/28/2007:SW:S:-:-	1.3090017109	3.8830708912

SW 3 Principal Component Scores	PC Size (No Rotation) normalized to 1	
EDA_Sample	PC 1	PC 2
USGS-07196090:6/4/2007:SW:S:-:-	1.3083552381	3.9769172813
USGS-07196090:7/12/2005:SW:S:-:-	1.3099523808	4.0174981513
USGS-07196090:7/3/2007:SW:S:-:-	1.4285174763	3.8227750432
USGS-07196090:8/23/2005:SW:S:-:-	1.3189053503	4.1609954515
USGS-07196090:8/6/2007:SW:S:-:-	1.3937346857	4.0996423434
USGS-07196090:9/16/2005:SW:S:-:-	1.3137563652	4.4216463441
USGS-07196500:1/9/2008:SW:S:-:-	1.4332966442	3.5216745386
USGS-07196500:10/3/2006:SW:S:-:-	1.3105059702	4.0468102797
USGS-07196500:11/17/2006:SW:S:-:-	1.3975528970	4.2366057681
USGS-07196500:12/13/2007:SW:S:-:-	1.3630498068	4.1318589303
USGS-07196500:12/5/2006:SW:S:-:-	1.3344613811	3.8617907332
USGS-07196500:12/6/2007:SW:S:-:-	1.2705144174	4.2483004546
USGS-07196500:2/12/2007:SW:S:-:-	1.3154529975	3.9047808237
USGS-07196500:3/11/2006:SW:S:-:-	1.3133655694	4.2976305880
USGS-07196500:4/13/2006:SW:S:-:-	1.2993609861	4.2776458352
USGS-07196500:4/27/2006:SW:S:-:-	1.2917134136	4.3553328264
USGS-07196500:4/27/2007:SW:S:-:-	1.2968709771	4.0595634423
USGS-07196500:4/29/2006:SW:S:-:-	1.4018415549	4.3366634865
USGS-07196500:4/3/2007:SW:S:-:-	1.3047822359	3.9813260957
USGS-07196500:5/1/2006:SW:S:-:-	1.3643601585	4.2888515410
USGS-07196500:5/25/2005:SW:S:-:-	1.3202859195	4.0072069264
USGS-07196500:6/13/2006:SW:S:-:-	1.3088324690	4.0294816557
USGS-07196500:6/13/2007:SW:S:-:-	1.3345419190	3.6401240282
USGS-07196500:6/25/2007:SW:S:-:-	1.2998851912	3.8167812132
USGS-07196500:6/7/2007:SW:S:-:-	1.3175138895	3.8423257236
USGS-07196500:7/11/2005:SW:S:-:-	1.3204498590	3.9461398433
USGS-07196500:7/16/2007:SW:S:-:-	1.3698197544	3.7983337234
USGS-07196500:7/5/2007:SW:S:-:-	1.3372955303	3.7438712354
USGS-07196500:7/6/2007:SW:S:-:-	1.3802286113	3.5908267184
USGS-07196500:8/24/2005:SW:S:-:-	1.3293554493	4.1056302680
USGS-07196500:9/19/2005:SW:S:-:-	1.3077305693	4.2192009364
USGS-07196500:9/19/2006:SW:S:-:-	1.4751993872	3.8788530885
USGS-07197000:1/10/2008:SW:S:-:-	1.2243743024	3.5430785850
USGS-07197000:1/25/2007:SW:S:-:-	1.2643215391	3.4421021759
USGS-07197000:10/4/2006:SW:S:-:-	1.2561909925	3.4711748042
USGS-07197000:10/4/2007:SW:S:-:-	1.2664764454	3.4955643744
USGS-07197000:10/9/2007:SW:S:-:-	1.2381077384	3.5084806643
USGS-07197000:11/17/2006:SW:S:-:-	1.3019789276	3.5150797858
USGS-07197000:12/13/2007:SW:S:-:-	1.2242553502	3.5227847926
USGS-07197000:12/4/2007:SW:S:-:-	1.2877982642	3.5444247297
USGS-07197000:12/5/2006:SW:S:-:-	1.2665738358	3.5082955476
USGS-07197000:2/6/2007:SW:S:-:-	1.2514287941	3.4724647772
USGS-07197000:4/19/2006:SW:S:-:-	1.2411197332	3.4394950088
USGS-07197000:4/29/2006:SW:S:-:-	1.3091525919	3.4270674217
USGS-07197000:4/4/2007:SW:S:-:-	1.2490730554	3.4773053590
USGS-07197000:5/1/2006:SW:S:-:-	1.2512965846	3.5184036373
USGS-07197000:5/11/2006:SW:S:-:-	1.3096354198	3.3564623539
USGS-07197000:5/4/2006:SW:S:-:-	3.2419367687	2.3855502512
USGS-07197000:6/12/2007:SW:S:-:-	1.2732544230	3.3750811683
USGS-07197000:6/14/2006:SW:S:-:-	1.2476063097	3.4338223679

SW 3 Principal Component Scores	**PC Size (No Rotation) normalized to 1**	
EDA_Sample	**PC 1**	**PC 2**
USGS-07197000:6/6/2007:SW:S:-:-	1.2540548240	3.4329831124
USGS-07197000:7/11/2005:SW:S:-:-	1.2836007498	3.4206661506
USGS-07197000:7/3/2007:SW:S:-:-	1.3680647380	3.3727993889
USGS-07197000:7/6/2007:SW:S:-:-	1.6409256033	3.2879024755
USGS-07197000:8/23/2005:SW:S:-:-	1.2901216540	3.4291351028
USGS-07197360:1/25/2007:SW:S:-:-	1.2335615339	3.3654722129
USGS-07197360:10/4/2006:SW:S:-:-	1.2348560026	3.5423728712
USGS-07197360:10/4/2007:SW:S:-:-	1.2210658357	3.5920187148
USGS-07197360:10/9/2007:SW:S:-:-	1.2202288216	3.6490906707
USGS-07197360:12/4/2006:SW:S:-:-	1.2278242495	3.5284396686
USGS-07197360:12/4/2007:SW:S:-:-	1.2550625840	3.6298115775
USGS-07197360:2/6/2007:SW:S:-:-	1.2158385182	3.3856312213
USGS-07197360:4/19/2006:SW:S:-:-	1.2198265429	3.6399895651
USGS-07197360:4/25/2007:SW:S:-:-	1.2450108674	3.5017422998
USGS-07197360:4/29/2006:SW:S:-:-	1.2684320063	3.5373524475
USGS-07197360:5/11/2006:SW:S:-:-	1.2536566039	3.4605617159
USGS-07197360:5/4/2006:SW:S:-:-	2.6007708694	2.7089987054
USGS-07197360:6/13/2006:SW:S:-:-	1.2238924458	3.5086058313
USGS-07197360:6/21/2007:SW:S:-:-	1.2548972664	3.4939242337
USGS-07197360:6/25/2007:SW:S:-:-	1.2284590505	3.5355041549
USGS-07197360:6/6/2007:SW:S:-:-	1.2275976879	3.5011082461
USGS-07197360:7/11/2005:SW:S:-:-	1.2615850890	3.4925523419
USGS-07197360:8/23/2005:SW:S:-:-	1.2643225824	3.5066033875